Molecular genetic analysis of secondary metabolite biosynthesis in cassava as an economic and nutritious plant

Bayoumi, Soad

Award date:
2008

Awarding institution:
University of Bath

Molecular Genetic Analysis of Secondary Metabolite Biosynthesis in Cassava as an Economic and Nutritious Plant

Soad Abdel Latief Hassan Bayoumi

A thesis submitted for the degree of Doctor of Philosophy

University of Bath

Department of Pharmacy and Pharmacology

Department of Biology and Biochemistry

October 2008

Table of Contents

Acknowledgments

I acknowledge the Almightiness of God who created all things including me and cassava.

I gratefully acknowledge the Egyptian Government for a fully-funded studentship.

My great appreciation is to my supervisors Dr Ian S. Blagbrough, Dr Michael G. Rowan and Dr John R. Beeching for their excellent supervision, advice, guidance and encouragement. I will be forever grateful for their advice. I have learned many techniques and had a good experience working with them.

I am grateful for accurate mass provision from Alison Smith and Christian Rehbein, University of Bath, and Bridget K. Stein, EPSRC National Mass Spectrometry Service Centre, University of Wales, Swansea. I thank Dr Timothy J. Woodman, Department of Pharmacy and Pharmacology, University of Bath, for his help with detailed NMR spectroscopy. I thank Dr Mary F Mahon, Department of Chemistry, University of Bath for the X-ray crystallographic analysis. I thank the research and technical staff in the Department of Pharmacy and Pharmacology: Kevin Smith, Jo Carter and Don Perry, and those in the Department of Biology and Biochemistry who take care of the cassava plants, Julia Watling and Angie Vale. Simon Bull provided support and help in the Biology lab. I am grateful to Dr Kim Reilly for all her help, support and friendship. Thanks also to people working at 5W 3.14 laboratory: Dr Nhum, Dr Osama, Dr Pilan, Dr Daniel, Dr Benjamin, Sawanya, Yiqiang, Hassan, Mostafa, Manal, Abdelkader, and to people working at 4S 1.52 laboratory: Mike Page a special thanks for his help, Glenda, Shazia, Ben, Vicky, Kate, Kimberly, Peter for the kind community working atmosphere. Thanks to Hanan for her friendship all along our way together in Bath.

Finally, but a big thank-you to my late father, Prof Abdel Latief Bayoumi, who got me interested in this whole research career, and I am aspiring to be like him. Many thanks to my husband Hosam, son El Hassan, mother Fatma and all the members of my family for their love and support.

Abstract

Cassava (*Manihot esculenta* Crantz Family Euphorbiaceae) is an important tropical food crop. However, harvested cassava roots have a shelf-life of only days due to post-harvest physiological deterioration (PPD). Within 1-3 days of harvesting, the roots show blue-black vascular streaking and are unpalatable. PPD includes altered gene expression and the accumulation of hydroxycoumarin secondary metabolites, e.g. scopoletin and esculetin, and their respective glucosides scopolin and esculin. In this research several important aspects of the biosynthesis of these phytochemically important hydroxycoumarins were resolved.

Stable isotopically labelled intermediates on the postulated biosynthetic pathways of scopoletin were fed to cassava cubes and PPD was allowed to occur. Ethanolic extracts of these deteriorated roots were separated (HPLC) and analysed (HRESI-MS). Incorporation (in both scopoletin and scopolin) of only 3 deuterons from *E*-cinnamic-2,3,2',3',4',5',6'-d_7 and *E*-cinnamic-3,2',3',4',5',6'-d_6 is strong support that the *E*-*Z*-isomerisation step is enzymatic and not photochemical. There are three hypothetical pathways for the biosynthesis of scopoletin via: 2',4'-dihydroxycinnamate, caffeate, or ferulate. High incorporation of label from *p*-coumaric-2-^{13}C, caffeic-2-^{13}C and ferulic-2-^{13}C acids was observed into labelled scopoletin and scopolin while there was only a small incorporation from ^{18}O-umbelliferone and ^{18}O-esculetin. We conclude that the major biosynthetic pathway to scopoletin and scopolin is via ferulic acid.

C$^{18}O_2$-enrichment of *E*-cinnamic and ferulic acids and feeding gave scopoletin containing only one ^{18}O-labelled oxygen atom. Therefore the lactonisation step is through *o*-hydroxylation and not via a postulated spirolactone-dienone intermediate. These results were confirmed by feeding experiments in an atmosphere of $^{18}O_2$-air which showed that the major isotopic peak was $^{18}O_3$-enriched scopoletin.

Three glucosyltransferases were isolated and identified from a cassava PPD-related cDNA library. These genes are expressed in the cassava storage root during PPD and they are also expressed in the fresh root. While one of these glucosyltransferases was novel, two had previously been isolated from cassava cotyledons.

List of Abbreviations

^{o}C	degree Celsius
ACh	acetylcholine
ACMD	African cassava mosaic disease
ACMV	African cassava mosaic virus
ALP	alkaline phosphatase
ALT	alanine aminotransferases
aq.	aqueous
AST	aspartate aminotransferases
bp	base pair
Calcl'd	calculated
CAZy	carbohydrate-active enzymes database
CBB	cassava bacterial blight
C3'H	*p*-coumaroyl shikimate/quinate 3'-hydroxylase
C4'H	cinnamate-4'-hydroxylase
CCoAOMT	caffeoyl-CoA *O*-methyltransferase
4'CL	4'-hydroxycinnamoyl CoA ligase
CCoAOMT	caffeoyl-CoA *O*-methyltransferase
COMT	caffeic acid methyltransferase
COSY	correlation spectroscopy
CM	CIAT *Manihot*
CIAT	Centro Internacional de Agricultura Tropical
CYP	cytochrome P450 monoxygenase
DEPC	diethyl pyrocarbonate
DEPT	distortionless enhancement by polarisation transfer
DMSO	dimethylsulfoxide
DOPA	3,4-dihydroxyphenylalanine
DPPH	2,2-diphenyl-1-picryhydrazyl hydrate
EDTA	ethylene diamine tetra acetic acid
ERK	extracelluar signal-regulated kinase

GADPH	glyceraldehyde-3-phosphate dehydrogenase
GSH	glutathione-*S*-transferase
GSH-RD	glutathione reductase
GTs	glycosyltransferases
HCT	shikimate/quinate hydroxycinnamoyltransferase
HPLC	high performance liquid chromatography
HMQC	heteronuclear multiple-quantum correlation
HR ESI MS	high resolution electrospray ionisation mass spectrometry
HR FABMS	high resolution fast atom bombardment mass spectrometry
LPL	lipoprotein lipase
MCOL	*Manihot* Colombia
MNGA	*Manihot* Nigeria
m.p.	melting point
m/z	mass over charge
NCBI	The National Center for Biotechnology Information databases
NIH	National Institute of Health
NMR	nuclear magnetic resonance spectroscopy
NOESY	nurclear overhouse effect spectroscopy
OD	optical density
PAL	phenylalanine ammonia lyase
PGF2α	prostaglandin $F_{2\alpha}$
PCR	polymerase chain reaction
PKC	protein kinase C
PPD	post-harvest physiological deterioration
ppm	part per million
PSPG	plant secondary product glycosyltransferases box
PVY	potato virus Y
RCF	relative centrifugal force
R_f	retention factor
R.H.	relative humidity
RNase	ribonuclease

RT	reverse transcriptase
SAR	structure-activity relationship
SDS	sodium dodecyl sulphate
s.d.	standard deviation
SDW	sterile distilled water
SOD	superoxide dismutase
SSC	salt sodium citrate
TAE	tris-acetate EDTA
TBE	tris-borate EDTA
TLC	thin layer chromatography
TMS	tetramethylsilane
TMV	tobacco mosaic virus
TOGT	tobacco glucosylransferase
Tris	tris (hydroxymethyl) methylamine
UDP-glucose	uridine diphosphate-glucose
UV	ultraviolet spectroscopy
UGT	UDP glycosyltransferase
X-Gal	5-bromo-4-chloro-3-indoyl-β-D-galactoside

Chapter 1
Introduction

1.1. Introduction and literature review

Despite sometimes being stigmatised as a "third world crop", cassava has been the focus of research in a few key areas due to the crop's importance as a food of the poor and its potential as a commodity in the wider economy.

Cassava (*Manihot esculenta* Crantz) is a perennial shrub (Fig. 1.1) belonging to the family Euphorbiaceae. The genus *Manihot* comprises 98 species of which *M. esculenta* is the most widely cultivated member (Rogers and Appan, 1973; Nassar et al., 2008). It originated in South America and subsequently has been distributed to topical and subtropical regions of Africa and Asia. The tuberous root of cassava is the fourth most important food source in the tropics after rice, maize, and sugar cane; it is a staple food for more than 500 million people due to its high starch content.

Fig. 1.1. Cassava shrub showing the edible tubers.

The cassava crop possesses an outstanding ability to grow in unfavourable environments: poor soil fertility and even extended periods (months) of drought, and this ability enhances its importance (Nassar, 2001). However, as an economic plant there are some problems that limit its utilization including:

Diseases

Cassava is susceptible to diseases such as cassava bacterial blight (CBB) which is caused by *Xanthomonas axonopodis pv manihotis* and African cassava mosaic disease (ACMD) caused by African cassava mosaic virus (ACMV) (Fregene et al., 2004; Kemp et al., 2004; Tomkins et al., 2004; Legg and Fauquet, 2004; Chellappan et al., 2004; Verdier et al., 2004; Taylor et al., 2004; Gómez-Vásquez et al., 2004; Kemp et al., 2005; Lopez et al., 2005; Zhang et al., 2005). The crop is also susceptible to viruses of the genus *Begomovirus* (family *Geminiviridae*) that cause cassava mosaic disease (CMD) (Bull et al., 2006).

Cyanogenic glycosides

Cassava contains cyanogenic glycosides, mainly linamarin and lotaustralin, at concentrations ranging from 6-370 mg/kg, which can release potentially toxic amounts of HCN upon rupturing of the cells (Rawel and Kroll, 2003; Charles et al., 2005). However, hydrogen cyanide can be efficiently removed from the root by various processing procedures (sun drying, heap fermentation, processing methods used to produce farinha in Brazil and gari in West Africa), but residual cyanogens in cassava foods may cause neurological disorders or paralysis, particularly in nutritionally compromised individuals (Cardoso et al., 2005). To address this problem transgenic cassava lines have been generated in which cyanogenic glycoside synthesis has been selectively inhibited (Siritunga and Sayre, 2004; Taylor et al., 2004; Jorgensen et al., 2005; Bradbury, 2006; Onyesom and Okoh, 2006) or by producing cultivars that promote rapid cyanide volatilization by over expression of hydroxynitrile lyase in roots leading to an elevation of the rate of cyanogen turnover (Siritunga and Sayre, 2007).

Low protein content

Cassava roots are perceived as having relatively low content of protein, vitamins and minerals compared to other sources of energy (such as maize and wheat). Protein content (Ceballos et al., 2006) as well as, vitamins and minerals (Welch, 2002; Bouis, 2003) can be increased by breeding methods. Also several biotechnological strategies have been discussed that might be used to increase protein level in transgenic cassava (Stupak et al., 2006).

Postharvest physiological deterioration (PPD)

Harvested cassava roots have a short shelf-life from two to three days due to post-harvest physiological deterioration (PPD), which starts as a black-blue to black vascular discoloration (vascular streaking) and then spreads to the parenchyma, thereby rendering the root unpalatable and unmarketable. PPD is followed by microbial deterioration. This rapid deterioration affects the economical value of the crop. For this reason, it is essential to understand PPD in order that it can be controlled, and this is the ultimate aim in this project.

It has been shown that within 15 min of the root being injured an oxidative burst occurs, which is followed by changes in the expression of genes, notably for catalase and peroxidase, related to the modulation of reactive oxygen species, and the accumulation of secondary metabolites, some of which show antioxidant properties. The interactions between these enzymes and compounds, in particular peroxidase and the coumarin, scopoletin, are largely confined to the vascular tissues where the visible symptoms of deterioration are observed (Reilly et al., 2003; Reilly et al., 2004). A superoxide dismutase (SOD) cDNA clone, mSOD2, encoding cytosolic copper/zinc SOD (CuZnSOD) was isolated from a PPD-related cDNA library, and its expression investigated in relation to environmental stress. mSOD2 gene in excised cassava leaves was highly induced by several abiotic stresses (Shin et al., 2005). More recently, cDNA microarray technology was used to characterise those genes that show significant change in expression during the PPD response. Seventy two non-redundant expressed sequence tags were identified which showed altered regulation during the post-harvest period. Of these 63 were induced, whilst 9 were down-regulated. Many of the up-regulated and

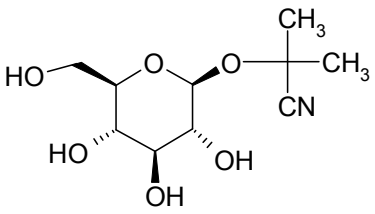
PPD-specific expressed sequence tags were predicted to play a role in cellular processes including reactive oxygen species turnover, cell wall repair, programmed cell death, ion, water or metabolite transport, signal transduction or perception, stress response, metabolism and biosynthesis, and activation of protein synthesis (Reilly et al., 2007).

There was correlation between both the dry matter and total carotenoid content and PPD in the roots of 101 cassava clones. PPD was positively, but weakly associated with dry matter content, and inversely associated with the total carotenoid content in roots. In addition, total carotenoid content and colour intensity were strongly and positively associated, suggesting that simple screening based on visual scoring of colour is adequate for initial selection of the roots of cassava clones with a relatively high total carotenoid content (Sanchez et al., 2006).

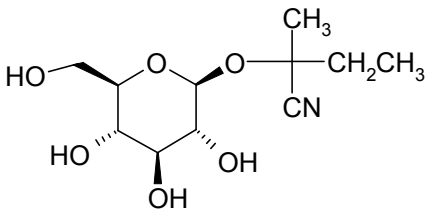
Compounds identified from cassava leaves and fresh roots

Cyanogenic glycosides

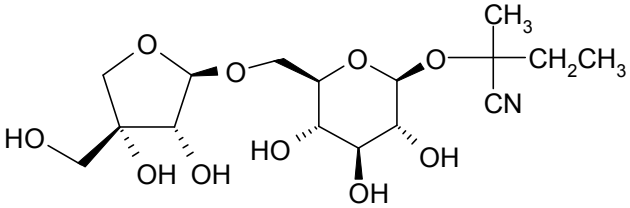
Linamarin and lotaustralin have been reported from cassava root and leaves (King and Bradbury, 1995; Du et al., 1995; Prawat et al., 1995) and 2-((6-*O*-(β-D-apiofuranosyl)β-D-glucopyranosyl)oxy)-2-methyl butanenitrile from cassava root (Prawat et al., 1995) (Fig. 1.2).



Linamarin



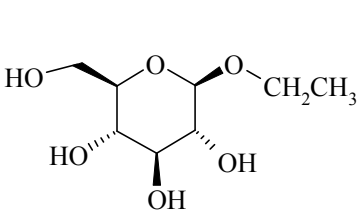
Lotaustralin



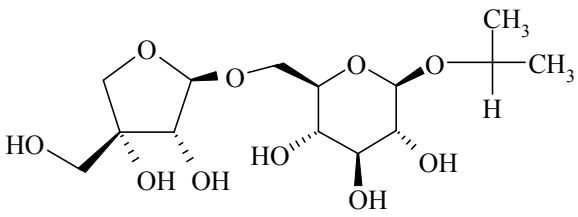
2-((6-*O*-(β-D-Apiofuranosyl)β-D-glucopyranosyl)oxy)-2-methylbutanenitrile

Fig. 1.2. Cyanogenic glycosides in cassava roots

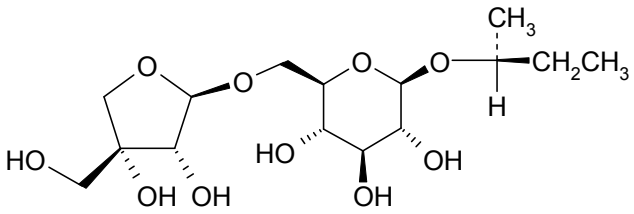
The following three novel glycosides of the corresponding alcohols were also found (Fig. 1.3) (King and Bradbury, 1995; Prawat et al., 1995).



Ethyl β-D-glucopyranoside



2-((6-*O*-(β-D-Apiofuranosyl)β-D-glucopyranosyl)oxy)propane

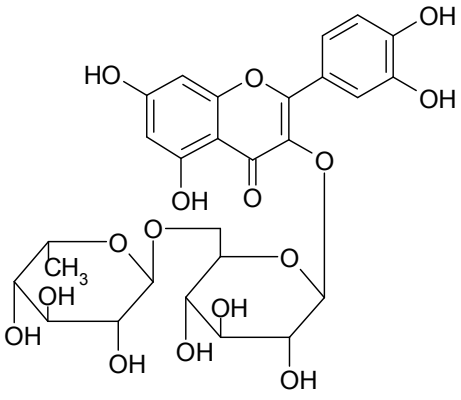


(2*S*)-((6-*O*-(β-D-Apiofuranosyl)β-D-glucopyranosyl)oxy)butane

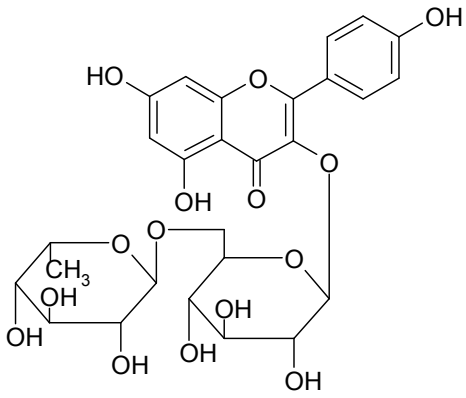
Fig. 1.3. Glycosides in cassava roots

Flavone-3-O-glycosides from cassava leaves

Rutin and Kamferol-3-O-rutinoside (Fig. 1.4) were identified in cassava leaves (Prawat et al., 1995).



Rutin (Quercetin-3-*O*-rutinoside)



Kampferol-3-*O*-rutinoside

Fig. 1.4. Flavone-3-*O*-glycosides from cassava leaves

Compounds identified from deteriorated cassava

Flavan-3-ols from deteriorated cassava root after 4-6 days

Three flavanols (Fig. 1.5) were identified and quantitatively compared in different cassava cultivars over a long period (over 7 days) of storage. These three flavanols ((+)-gallocatechin, (+)-catechin and (+)-catechin gallate) are not synthesized in healthy root, but are detectable in low concentrations during the first 24-48 h after root harvesting. After 4-6 days there was some accumulation of all three flavanols, after 7 days there was a rapid decline of all flavanols. However, these flavanols are potential antioxidants and the authors could not relate their accumulation in cassava roots to early storage disorders or wound responses because they start to accumulate only after 4-6 days (Buschmann et al., 2000a).

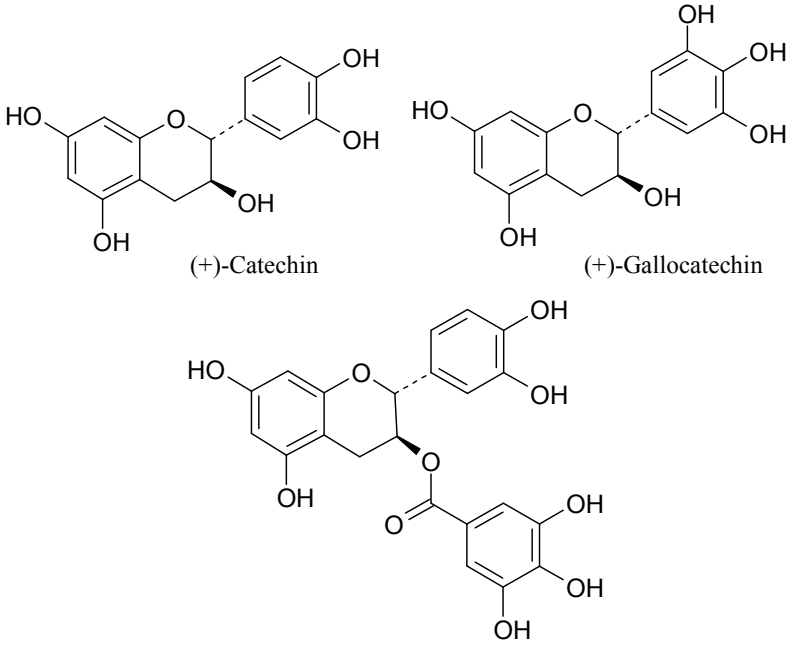


Fig. 1.5. Flavan-3-ols from deteriorated cassava root.

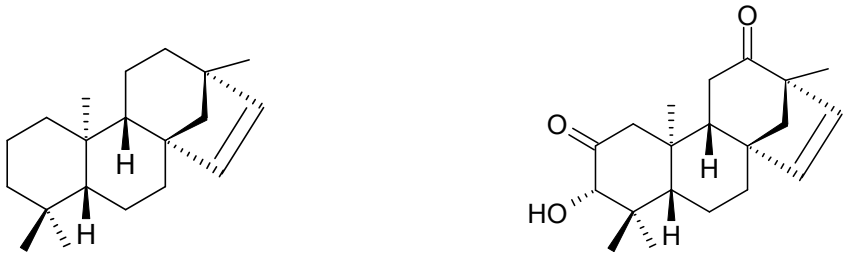
Diterpene compounds from deteriorated cassava root after 7 days

Twenty-two diterpenic stress metabolites, most of which were novel, were isolated and identified from 5 kg of cassava root tissues damaged by cutting or fungal-infection (Sakai and Nakagawa, 1988). The metabolites can be classified into the following four skeletal types (number of components shown): ent-beyerane (10

components), ent-pimarane (9 components), ent-atisane (2 components), and ent-kaurane (1 component). Diterpenes are not common plant stress metabolites.

ent-Beyerane

This was the largest group of diterpenic stress metabolites and yucalexin B-1((+)-stachene) was the mother skeleton of this family. Yucalexin B-9 (ent-3β-hydroxybeyer-15-ene-2,12-dione) was the most abundant diterpenoid component in deteriorated cassava root (47.8 mg per 5 kg fresh weight) (Fig. 1.6).

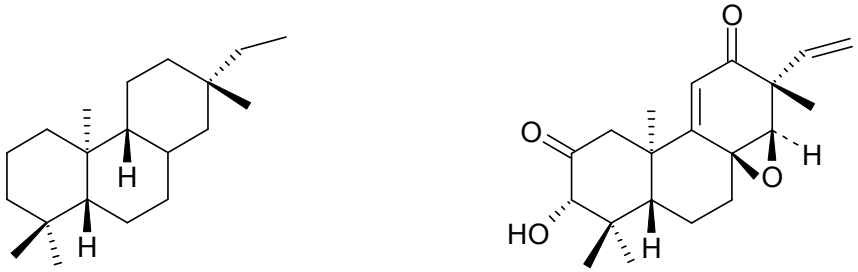


Yucalexin B-1 Yucalexin B-9

Fig. 1.6. ent-Beyerane from deteriorated cassava root

ent-Pimarane

The novel compound, yucalexin P-8 (ent-8α,14α-epoxy-3β-hydroxypimara-9(11),15-diene-2,12-dione) was the second most abundant diterpenoid component in the cassava root (18 mg per 5 kg fresh weight) (Fig. 1.7).

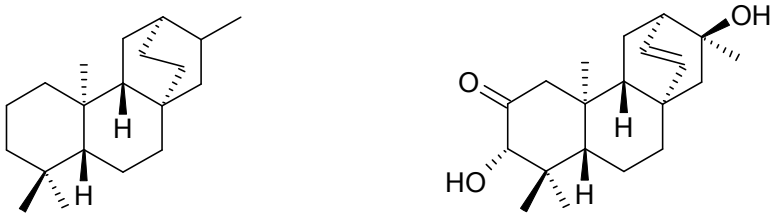


ent-Pimarane Yucalexin P-8

Fig. 1.7. ent-Pimarane from deteriorated cassava root

ent-Atisane

Yucalexin A-19 (ent-3β,16α-dihydroxyatis-13-en-2-one) was also a novel compound. It was the third most abundant diterpenoid component in the cassava root (7.3 mg per 5 kg fresh weight) (Fig. 1.8).

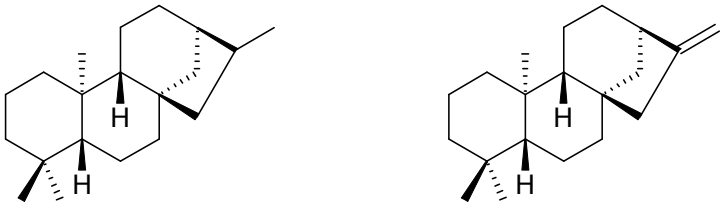


ent-Atisane Yucalexin A-19

Fig. 1.8. ent-Atisane from deteriorated cassava root

ent-Kaurane

ent-Kaurene (Fig. 1.9) was the sole component of this type (Sakai et al., 1986).



ent-Kaurane Yucalexin K-3 (ent-Kaurene)

Fig. 1.9. ent-Kaurane from deteriorated cassava root

Steroids from three deteriorated cassava root cultivars

Steroids (Fig. 1.10) were identified from physiologically and microbially damaged roots of cassava cultivars Golden yellow, Hawaiian, and Okinawa (Sakai et al., 1986) in comparison with healthy roots. Cassava roots were cut into small pieces and air dried for two days in room conditions (healthy roots) while other slices were incubated for 2, 3, and 7 days then air dried for two days in room conditions (physiologically damaged roots); also roots were incubated for four days with *Botryodiplodia theobromae* (microbially damaged root). Healthy cassava roots contained the widespread plant sterols, β-sitosterol, campesterol, stigmasterol and cholesterol. In addition to these

common sterols, both physiologically and microbially damaged roots contained an interesting series of conjugated enones based on the same skeletons.

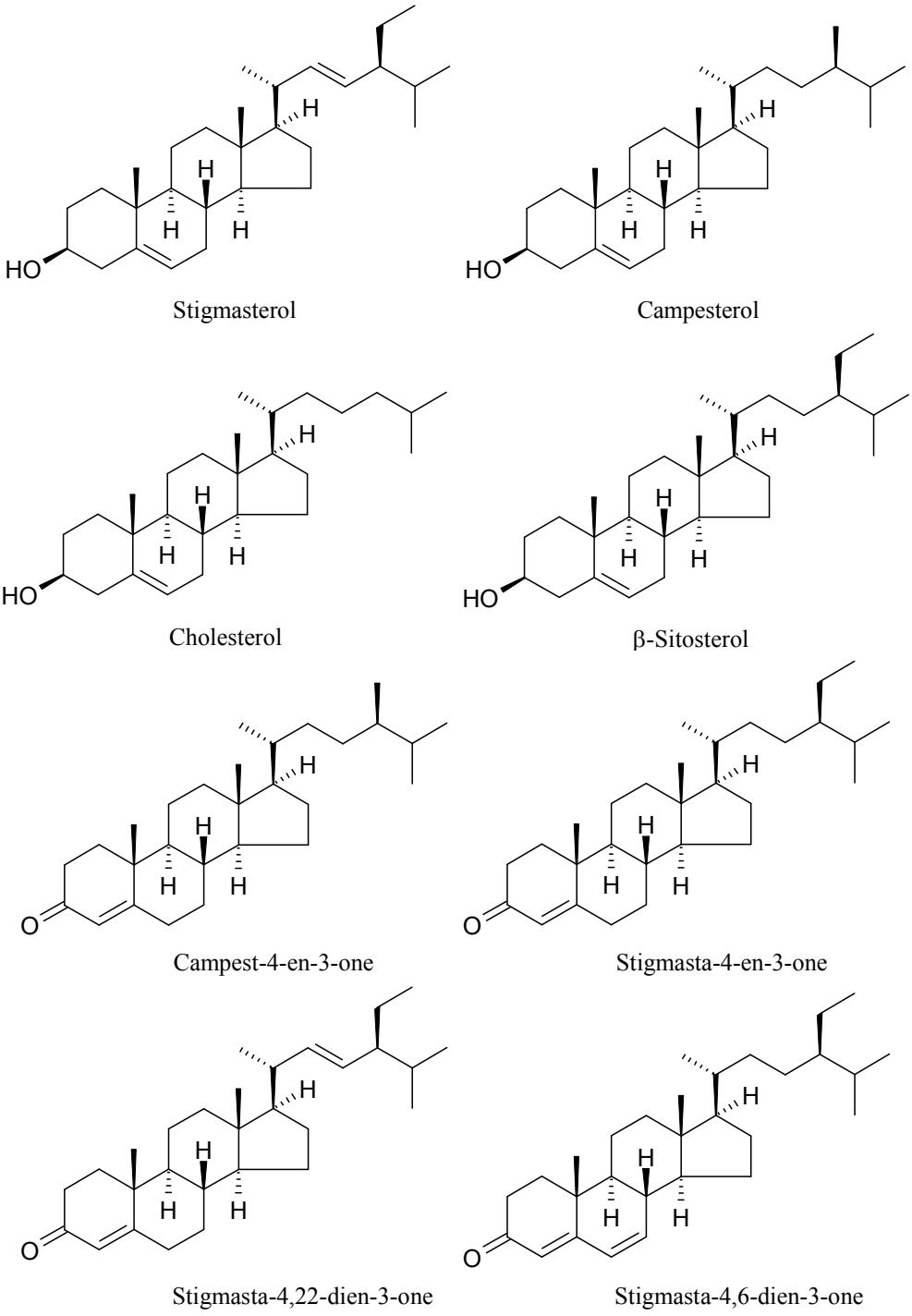


Fig. 1.10. Steroids from healthy and deteriorated cassava roots.

Fatty acids and esters

Palmitic, oleic and linoleic acids in addition to their methyl esters and mono glycerides were found in healthy cassava roots as well as in physiologically and microbially damaged roots (Sakai et al., 1986).

Hydroxycoumarins from deteriorated cassava root

Hydroxycoumarins are found as secondary metabolites in the plant kingdom; some members of the plant families Rutaceae, Solanaceae and Apiaceae accumulate large amounts. In cassava roots within 24-48 h after harvesting, there is a rapid accumulation of hydroxycoumarins which were identified as scopoletin, scopolin (its glucoside) and esculetin, esculin (its glucoside) (Fig. 1.11) and, due to the 150-200 fold increase of scopoletin concentration, it was assumed that scopoletin is involved in PPD (Tanaka et al., 1983; Wheatley and Schwabe, 1985; Buschmann et al., 2000b).

Scopoletin accumulation tended to be more pronounced in cassava cultivars with high susceptibility to deterioration then, after 6 days, the scopoletin concentration was less pronounced due to the metabolism of scopoletin to an insoluble blue-black coloured product (vascular streaking of PPD) by means of a peroxidase (Reilly et al., 2003). This assumption was explained on the basis that peroxidase activity increased simultaneously with decrease in scopoletin concentration as well as marked increase in the blue-black colouration of the vascular and parenchymatous tissue (Buschmann et al., 2000b).

Hydroxycoumarins are derived from L-phenylalanine with the action of phenylalanine ammonia lyase (PAL), and so via *E*-cinnamic acid (Fig. 1.11) in the phenylpropanoid pathway (Strack, 1997; Petersen et al., 1999). Three hypothetical pathways for the biosynthesis of scopoletin (Kai et al., 2006) have been proposed from studies in various plant species. These pathways are via: 2',4'-dihydroxycinnamate, 3',4'-dihydroxycinnamate (caffeate), or 4'-hydroxy-3'-methoxycinnamate (ferulate) (Fig. 1.12) and they will be investigated in cassava roots during PPD in Chapter 4.

Enzymes catalyse these biosynthetic reactions in the three proposed pathways for the biosynthesis of scopoletin and scopolin in different plants e.g. PAL, cinnamate-4'-hydroxylase, *p*-coumaroyl shikimate/quinate 3'-hydroxylase, caffeoyl-CoA *O*-methyltransferase and scopoletin glucosyltransferase (GT) (Fig. 1.12). These enzymes

will be discussed later in Chapter 5. The *E-Z* isomerisation of cinnamic acids and the *o*-hydroxylation step are investigated and discussed in detail in Chapters 3 and 4.

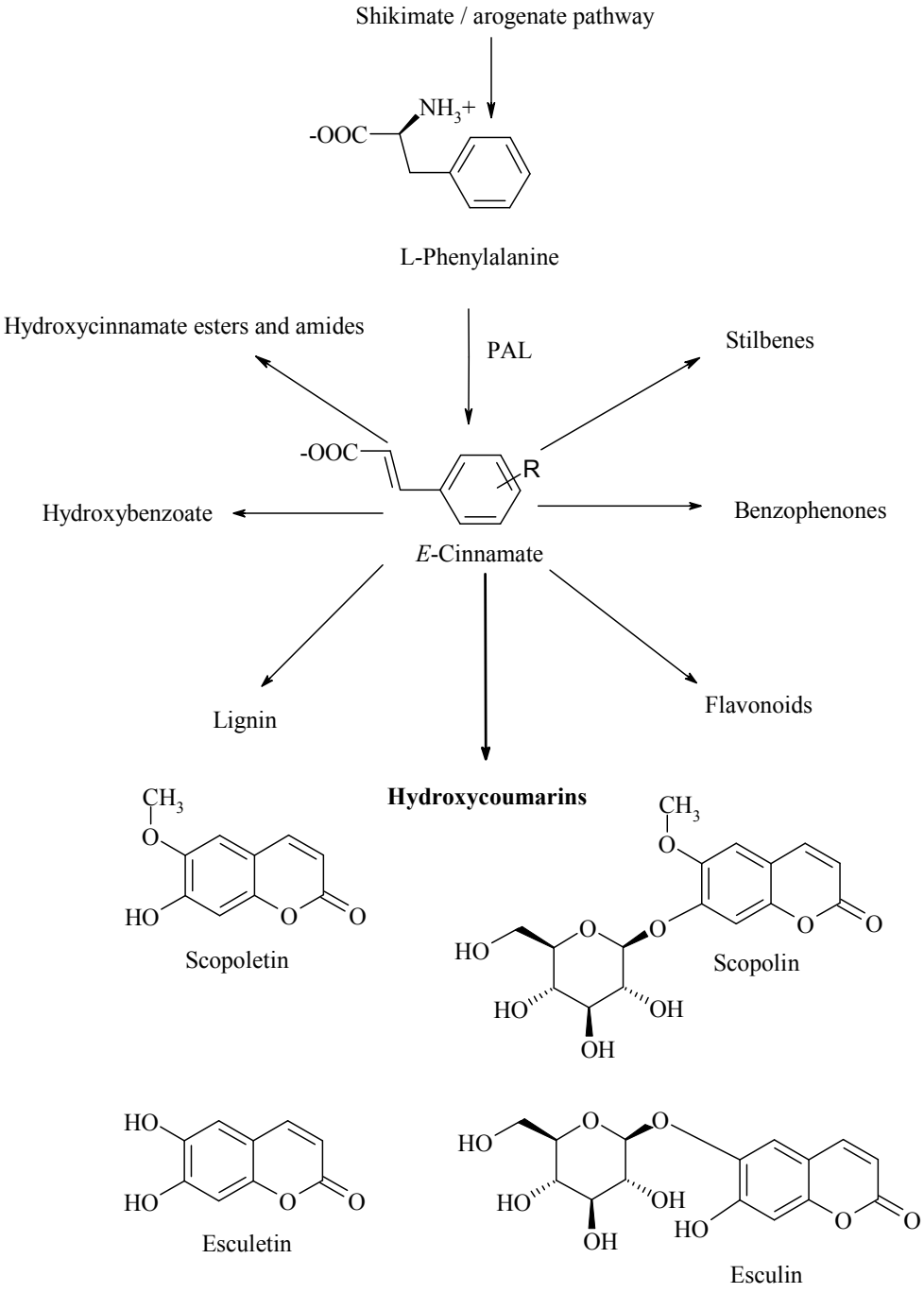


Fig. 1.11. Biosynthetic pathways of phenylpropanoids and hydroxycoumarins.

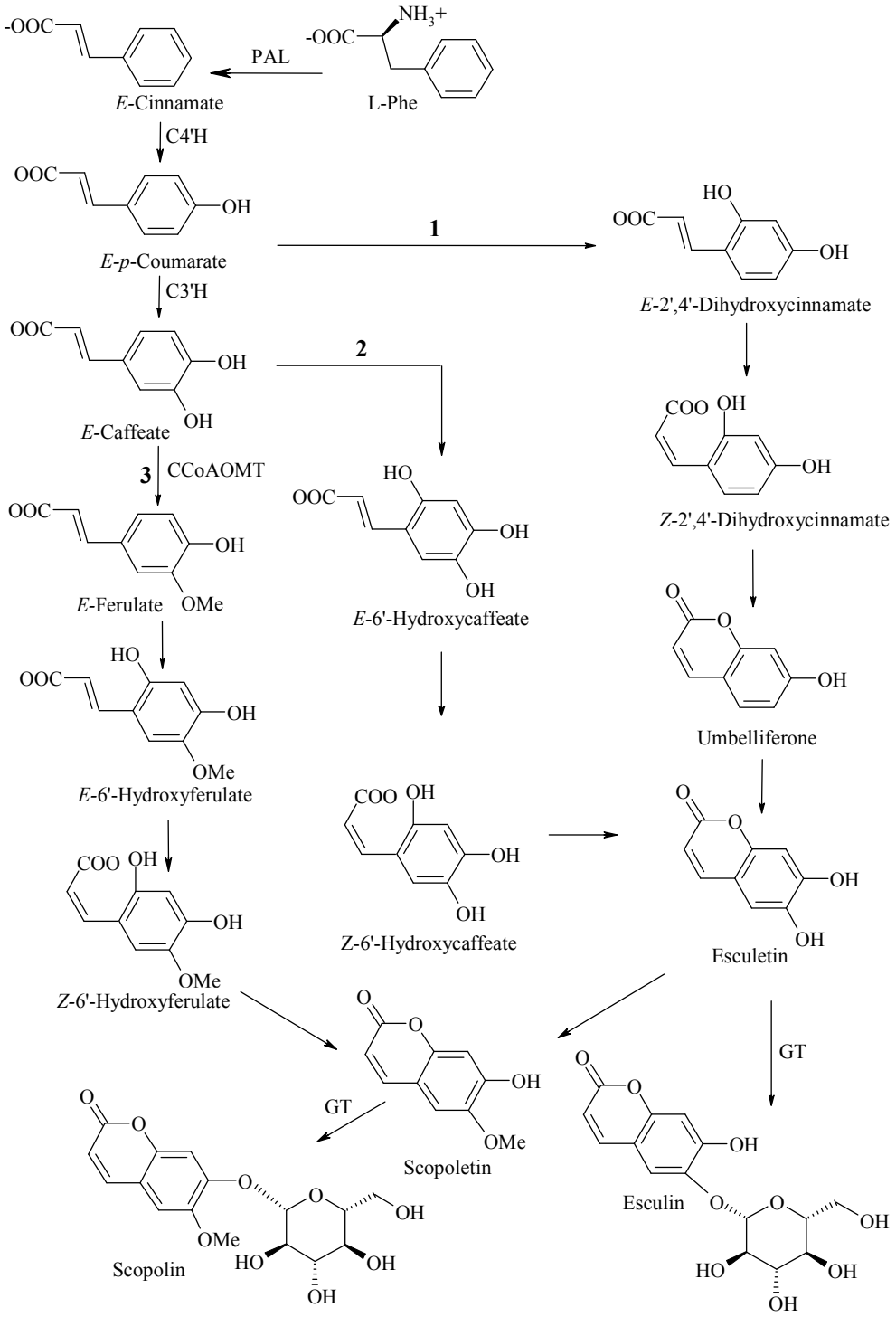


Fig. 1.12. Three proposed metabolic pathways in the biosynthesis of scopoletin (for clarity the carboxylate anion has been omitted after *E*-cinnamate). Abbreviations: caffeoyl-CoA *O*-methyltransferase (CCoAOMT), cinnamate-4'-hydroxylase (C4'H), *p*-coumaroyl shikimate/quinate 3'-hydroxylase (C3'H), glucosyltransferase (GT) and phenylalanine ammonia lyase (PAL).

The importance of hydroxycoumarins for plants

Hydroxycoumarins have different bioactivities and contribute essentially to the persistence of plants, being involved in processes such as defence against phytopathogens, response to abiotic stresses, regulation of oxidative stress, and probably hormonal regulation (Bourgaud et al., 2006).

Coumarins inhibit growth regulators such as indole acetic acid, cytokinin and gibberellins to regulate plant growth (Brown, 1981). Hydroxycoumarins are normal constituents in some plants but scopoletin in sunflower and other plants accumulates following mechanical wounding, insect feeding damage and fungal or bacterial infection, and may function as phytoalexins (Olson and Roseland, 1991; Roseland et al., 1991).

The high disease resistance of the amphidiploid hybrid of *Nicotiana glutinosa x Nicotiana debneyi* was associated with high constitutive levels of scopoletin and scopolin. Scopoletin showed a direct antimicrobial activity against *Cercospora nicotianae*, *Phytophthora parasitica* var. *nicotianae*, *Pseudomonas syringae* pvs. *tabaci* and *syringae* and tobacco mosaic virus when added to synthetic growth media, mixed with the inoculum or sprayed onto tobacco plants prior to inoculation. The high amount of toxic phenolics in the leaves of this hybrid coincided with its high disease resistance (Goy et al., 1993).

Scopoletin as a phytoalexin was induced by various fungi in leaves of *Hevea brasiliensis*. Scopoletin showed antifungal activity and there was a correlation between scopoletin induction and the resistance of rubber leaves to the infection with *Microcyclus ulei* (Giesemann et al., 1986). *Hevea* leaflets inoculated with *M. ulei* produced scopoletin, in the infected zones. The clones with total resistance and those with a high level of partial resistance displayed an early accumulation of scopoletin that intensified and lasted longer than 48 h in the fully resistant clones. The fungitoxic effect of scopoletin was verified *in vitro* on *M. ulei* where 2 mM concentrations were sufficient to strongly inhibit germ tube elongation and conidium germination. *In situ*, 24 h after inoculation, conidium germination and number of infection sites were lower in the most resistant clones. The fungitoxic effect of scopoletin was tested on two other leaf pathogens of the rubber-tree, *Colletotrichum gloeosporioides* and *Corynespora*

cassiicola. Concentrations double or more than those tested on *M. ulei* were required for total inhibition of germination and germ tube elongation (Garcia et al., 1995).

Inoculation of cell suspension cultures derived from *Ulmus pumila* with spores of *Ophiostoma ulmi* induced the accumulation of scopoletin. In *in vitro* bioassays, scopoletin showed a direct antifungal activity against *O. ulmi*, spore germination being more sensitive to this inhibitory activity than mycelial growth (Valle et al., 1997).

Antifungal responses elicited by yeast in cassava suspension cells and leaves resulted in anionic forms of peroxidase which oxidized scopoletin, with highest activity by isoform, pl 3.6, present in all samples. Unidentified phenolics and possibly scopolin increased post-elicitation, but there was no enhancement of scopoletin, rutin or kaempferol-3-*O*-rutinoside concentration. Fungitoxicity was determined against the cassava pathogens *Fusarium solani*, *F. oxysporum* and the saprotroph *Trichoderma harzianum*, fungal germ tube elongation was inhibited more than germination by esculetin, ferulic acid, quercetin and scopoletin. *T. harzianum* was generally more sensitive than the pathogens and was inhibited by 50 µg/ml of ferulic acid and quercetin, and 10 µg/ml of scopoletin. Phenolic levels in cells were, theoretically, too low to be inhibitory. However, in combination and when oxidized, they may contribute to defence, because oxidation of esculetin and scopoletin by peroxidase and of esculetin by tyrosinase enhanced their fungitoxicity up to 20-fold (Gómez-Vásquez et al., 2004).

Transgenic tobacco plants over-expressing a salicylate- and pathogen-inducible glucosyltransferase (TOGT) showed enhanced resistance against infection from potato virus Y (PVY). Under non-infectious conditions, profiles of phenylpropanoids in leaves of transgenic lines were similar to those of controls. Feeding experiments with leaf-discs demonstrated a higher capacity for glucosylation of scopoletin. After inoculation with PVY the transgenic lines showed similar formation of necrotic leaf lesions, but significantly decreased levels of virus coat-protein when compared with control plants. This, implies that the activity of TOGT and the subsequent accumulation of glucosylated coumarins represent an important step in the cascade of events resulting in confinement of viral pathogens (Matros and Mock, 2004).

TOGT-depleted tobacco plants were generated by antisense expression. After inoculation with tobacco mosaic virus (TMV), TOGT-inhibited plants exhibited a

significant decrease in scopolin and a decrease in TOGT activity. Unexpectedly, free scopoletin levels also were reduced in TOGT antisense lines. Scopolin and scopoletin reduction in TOGT-depleted lines resulted in a strong decrease of the blue fluorescence in cells surrounding TMV lesions and was associated with weakened resistance to infection with TMV providing evidence of the crucial role of glycosyltransferases in plant defence responses. It was proposed that TOGT-mediated glucosylation is required for scopoletin accumulation in cells surrounding TMV lesions, where this compound could exert a direct antiviral effect (Chong et al., 2002).

Scopoletin and scopolin were tested *in vitro* for their biological activity (concentration-activity relationships) using several bioassays: germination of proso-millet (*Panicum milliaceum* L.) seed; mycelial growth of the sweet potato fungal pathogens *Fusarium oxysporum* Schlecht. *F. batatas* (Wollemv.) Snyd. & Hans, *F. solani* (Sace.) Mart., *Lasiodiplodia theobromae* (Pat.) Griffon & Maubl., and *Rhizopus stolonifer* (Ehr. ex Fr.) Lind; and growth and mortality of diamondback moth [*Plutella xylostella* (L.)] larvae raised on an artificial diet. The glycoside scopolin showed little activity apart from moderate inhibition of *F. oxysporum*. The aglycone scopoletin inhibited seed germination and larval growth; however, at much higher concentrations than were measured in the tissues. Mycelial growth of the four pathogenic fungi, however, was inhibited at concentrations occurring in some sweet potato clones (Peterson et al., 2003).

Chlorogenic acid, scopolin, scopoletin, and structurally related compounds were evaluated for inhibition of growth of *Phytophthora parasitica* var. nicotianae (black shank) fungus in a laboratory bioassay. Scopoletin gave 39% inhibition at 1000 ppm dosage. Scopoletin and esculetin were active in the laboratory bioassay, but their glucose-derivatives were found to be completely inactive (Snook et al., 1991).

Scopolin showed inhibitory activity to *Sclerotinia* at similar doses to scopoletin (fungitoxic effect), scopolin is known to be less phytotoxic than ayapin (6,7-methylenedioxycoumarin) and scopoletin, its accumulation may well confer head rot resistance with minimal plant damage and might be one of the bases for resistance to *Sclerotinia* (Prats et al., 2006).

Potential therapeutic activitiesof hydroxycoumarins

Ethnopharmacological studies of a wide variety of traditional medicinal plants coupled with activity directed fractionations have revealed many potential therapeutic applications, though few have been subjected to detailed pharmacological evaluation.

Antimicrobial activity

Sixteen coumarins, 7,8-dihydroxycoumarin, umbelliferone, scoparone (6,7-dimethoxycoumarin), esculetin , 6-hydroxy-7-methoxycoumarin, herniarin (7-methoxycoumarin), and scopoletin, 6,7-diacetoxy coumarin, 6-methoxy-7-acetylcoumarin, and 6-acetoxy-7-methoxycoumarin, 8-Methoxypsoralen, 8-acetyl-7-hydroxycoumarin, 7,8-dihydroxy-6-methoxycoumarin, 6,7-dimethoxy-4-methylcoumarin, 5,7-dihydroxy-4-methylcoumarin, 4-hydroxycoumarin, 4-hydroxy-6,7-dimethylcoumarin were tested against a panel of bacteria and fungi. Their antibacterial activity was determined on *Bacillus subtilis*, *Escherichia coli*, *Proteus mirabilis*, *Klebsiella pneumoniae*, *Salmonella typhi*, *Salmonella* sp., *Shigella boydii*, *Shigella* sp*., Enterobacter aerogenes, Enterobacter agglomerans*, *Sarcina lutea*, *Staphylococcus epidermidis*, *Staphylococcus aureus*, *Yersinia enterolitica*, and *Vibrio cholerae*. The evaluated fungi were *Aspergillus niger*, *Penicillium notatum*, *Fusarium moniliforme*, *Fusarium sporotrichum*, *Rhizoctonia solani*, and *Trichophyton mentagrophytes*. The most active compounds against Gram-positive and negative bacteria were the dihydroxylated coumarins, 7,8-dihydroxy-6-methoxycoumarin and 7,8-dihydroxycoumarin. In addition, 8-acetyl-7-hydroxycoumarin, 7,8-dihydroxy-6-methoxycoumarin and 7,8-dihydroxycoumarin, 6,7-dimethoxy-4-methylcoumarin, scoparone, and esculetin showed an interesting activity against *V. cholerae*, a key bacterium in contaminated water; 8-acetyl-7-hydroxycoumarin, 7,8-dihydroxy-6-methoxycoumarin and 7,8-dihydroxycoumarin were the most active. The dimethoxy compounds 6,7-dimethoxy-4-methylcoumarin and scoparone showed a strong activity against fungal strains, especially *T. mentagrophytes* and *R. solani* (Cespedes et al., 2006).

Esculetin and its diacetate exhibited a marked inhibitory effect on Newcastle disease virus replication in cell cultures at concentrations of 36 μM and 62 μM,

respectively. These compounds were selected from ten hydroxycoumarin derivatives through an *in vitro* antiviral screen involving viruses of the picorna-, orthomyxo-, paramyxo-, and herpes virus families (Galabov et al., 1996).

Scopoletin showed inhibitory activity against *Mycobacterium smegmatis* (Mativandlela et al., 2007), exhibiting a minimum inhibitory concentration of 7.8 mg/ml.

Antioxidant

The antioxidant activity of hydroxycoumarins, in particular esculetin and scopoletin can occur either through direct radical scavenging or by changing the activity of oxidative enzymes (peroxidase, cycloxygenase, lipoxygenase and xanthine oxidase). This activity might be responsible for the results in more complex biological test systems.

Free radical scavenging

Methanolic extracts of *Artemisia iwayomogi* showed excellent peroxynitrite ($ONOO^-$) scavenging activity in a dose-dependent manner. Apigenin 7-methylether (genkwanin), apigenin 7,4'-di-*O*-methylether, jaceosidin, and scopoletin from the EtOAc fraction and chlorogenic acid, 2,4-dihydroxy 6-methoxy acetophenone 4-*O*-β-D-glucoside, quebrachitol, and scopolin from the n-BuOH fraction were identified. The scavenging activities of chlorogenic acid (IC50 = 0.52 +/- 0.04 μM), genkwanin (IC50 = 1.01 +/- 0.10 μM), and scopoletin (IC50 = 1.03 +/- 0.15 μM) were higher than or comparable to a well-known $ONOO^-$ scavenger, penicillamine (IC50 = 1.76 +/- 0.18 μM), suggesting that those compounds might be developed as effective $ONOO^-$ scavengers for, prevention of the $ONOO^-$ involved diseases (Kim et al., 2004).

The antioxidant activity of *Artemisia montana* was determined by measuring the radical scavenging effect on 1,1-diphenyl-2-picrylhydrazyl (DPPH) radical and inhibitory activity against free radical generation of hepatocytes (AC(2)F). The methanol extract of *A. montana* showed strong radical scavenging activity at a concentration of 10.1 μg/ml. Esculetin and luteolin 7-*O*-rutinoside were isolated; the antioxidant activity of these compounds were comparable to that of L-ascorbic acid (Kim et al., 2000).

The chloroformic fraction of *Fraxinus rhynchophylla* (Cortex Fraxini) is rich in *o*-dihydroxy simple coumarins, esculetin and fraxetin which have better (or equal) radical-scavenging capacities than have flavonoids and known antioxidants, including ascorbic acid. This fraction exhibited high radical-scavenging potency; its antioxidant activity was correlated with its reducing power, but not iron chelation (Wu et al., 2007)

Eight coumarin derivatives were put under oxidative conditions. Free radicals was formed by 1,1-diphenyl-2-picrylhydrazyl hydrate and 5,5-dimethyl-1-pyrroline-N-oxide electron spin resonance. Esculetin was still the most potent agent at protecting cells against reactive oxygen species among other tested coumarins (Lin et al., 2008).

Inhibition of oxidant enzyme activity

Esculetin inhibits lipoxygenase activity in soybean in a concentration dependent manner. In the study comparing the inhibition abilities for lipoxygenase activity among selected antioxidants, esculetin was significantly better than butylated hydroxyanisole and α-tocopherol (Lee and Lillard, 1997). Lipoxygenase activity was inhibited by sodium azide, esculin, mercuric chloride and quercetin (Mohammadi and Karr, 2003).

Scopoletin inhibits horseradish peroxidase reactions with molecular oxygen and indole-3-acetic acid (growth regulator). Peroxidase activity was decreased by increasing scopoletin concentration. Scopoletin or related coumarins with peroxidase and the suppression of indole-3-acetic acid degradation may provide a possible control mechanism over the oxidative degradation of indole-3-acetate by this plant enzyme (Sirois and Miller, 1972).

Xanthine oxidase catalyses the oxidation of hypoxanthine to xanthine and can further catalyse the oxidation of xanthine to uric acid. Esculetin is a strong xanthine oxidase inhibitor; the 7-hydroxy coumarin group being essential for activity. The 6-hydroxy group present in esculetin enhanced the activity, whereas substitution by the 6-methoxy group, e.g. scopoletin, reduced the inhibitory effect. Furthermore, the 6-glycoside group in esculin strongly decreased the inhibitory effect. Scoparone, the fully methylated derivative of esculetin also showed low activity (Chang and Chiang, 1995). Esculetin had the highest affinity toward the binding site of xanthine oxidase, mainly due to the interaction of the 6-hydroxyl with the E802 residue of xanthine (oxidase

based molecular modeling). Esculetin can also block NADPH oxidase. It generates superoxide by transferring electrons from NADPH inside the cell across the membrane and coupling these to molecular oxygen to produce the superoxide, which is a highly-reactive free-radical (Lin et al., 2008).

Scopoletin isolated from a water-extract of *Artemisia feddei* showed inhibition of nitric oxide synthesis in a dose-dependent manner in murine macrophage-like RAW 264.7 cells stimulated with interferon-gamma and lipopolysaccharide. The inhibition of nitric oxide synthesis by scopoletin was due to suppression of inducible nitric oxide synthase mRNA and inducible nitric oxide synthase protein, as determined by Northern and Western blotting, respectively (Kang et al., 1999).

Anti-hepatotoxic

Esculetin was investigated in *Cichorium intybus* and *Bougainvillea spectabilis* for its possible protective effect against paracetamol and CCl_4-induced oxidative hepatic damage due to its antioxidant activity. Paracetamol produced 100% mortality at the dose of 1 g/kg in mice while pre-treatment of animals with esculetin (6 mg/kg) reduced the death rate to 40%. Oral administration of paracetamol (640 mg/kg) produced liver damage in rats as manifested by the rise in serum enzyme levels of alkaline phosphatase (ALP) and aminotransferases (AST and ALT). Pre-treatment of rats with esculetin (6 mg/kg) prevented the paracetamol-induced rise in serum enzymes. The hepatotoxic dose of CCl_4 (1.5 ml/kg; orally) also raised serum ALP, AST and ALT levels. The same dose of esculetin (6 mg/kg) was able to prevent the CCl_4-induced rise in serum enzymes. Esculetin also prevented CCl_4-induced prolongation in pentobarbital sleeping time confirming hepatoprotectivity. These results indicate that esculetin possesses anti-hepatotoxic activity and the presence of this compound in *Cichorium intybus* and *Bougainvillea spectabilis* may explain the folkloric use of these plants in liver damage (Gilani et al., 1998).

Anti-inflammatory activity

Esculetin, scopolin, and scopoletin were isolated from *Santolina oblongifolia* showed marked activity as inhibitors of eicosanoid-release from ionophore-stimulated mouse peritoneal macrophages (Silvan et al., 1996).

Five coumarins, psoralen (7*H*-furo[3,2-g]benzopyran-7-one), xanthotoxin (9-methoxyprosalen), scopoletin, umbelliferone, and (+)-marmesin (a furoanocoumarin), identified in the root of *Dystaenia takeshimana*, showed cyclooxygenase-2 and 5-lipoxygenase dual inhibitory activity by their effects on the production of prostaglandin D-2 and leukotriene C-4 in mouse bone marrow-derived mast cells. These results suggest that the anti-inflammatory activity of *D. takeshimana* might in part occur via the inhibition of the generation of eicosanoids (Kim et al., 2006b).

Scopoletin, isolated from an aqueous extract of *Senra incana*, inhibited prostaglandin synthetase in a dose-dependent way. Compared to aspirin, the potency of scopoletin was about five times higher. On topical application, scopoletin dose-dependently inhibited ethyl phenylpropiolate-induced edema of the rat ear. The active dose range was 1-10 μg/ear (Farah and Samuelsson, 1992). Scopoletin has anti-inflammatory activity (Moon et al., 2007) by inhibiting the induced production of inflammatory cytokines of phorbol 12-myristate 13-acetate and A23187 (ion carrier) and by regulating the IκB/NF-κB signal cascade. Scopoletin is a specific inhibitor of the production of inflammatory cytokines in mast cells, and this inhibition might explain its beneficial effect in the treatment of chronic inflammatory diseases.

Anticancer

A number of ethnopharmacological studies have suggested potential anticancer activities for hydroxycoumarins. The ethyl acetate extract of *Artemisia argyi* leaves showed substantial inhibition in a cell proliferation assay using human CCRF-CEM leukaemia cells. Bioassay-guided fractionation of the extract led to the isolation of scopoletin and isoscopoletin as the active principles (Adams et al., 2006). Aqueous extracts or infusions obtained from the flowers of *Tilia* species and scopoletin, the main component in the dichloromethane extract showed antiproliferative action on BW 5147 lymphoma cells (Arcos et al., 2006).

More detailed studies have examined structure activity relationships and pharmacological mechanisms. A recent review concluded that coumarin is a natural substance that has shown anti-tumour activity *in vivo*, with the effect believed to be due to its metabolites (e.g. 7-hydroxycoumarin). This review was based on a recent study has shown that 7-hydroxycoumarin inhibits the release of cyclin D1, which is overexpressed in many types of cancer. This knowledge may lead to its use in cancer therapy. Esculetin inhibits growth and cell cycle progression by inducing arrest of the G(1) phase in HL-60 leukaemia cells, resulting from the inhibition of retinoblastoma protein phosphorylation. The effects of coumarins and coumarin-related compounds were investigated on a panel of cell-lines. The most recent work involves two cell-lines, MCF-7 a breast carcinoma and A549 a lung carcinoma. Microtitre assays for cell growth were performed along with real-time analysis of cell viability. These studies suggested that both genistein and esculetin exerted the most potent inhibitory effect on cell growth in comparison to the other compounds (Lacy and O'Kennedy, 2004).

Twenty eight coumarins, including 7 furocoumarins, were examined for their activity of induction of terminal differentiation of human promyelocytic leukemia cells (HL-60) by nitro blue tetrazolium reducing, esterase and phagocytic activities. Esculetin, nordalbergin (6,7-dihydroxy-4-phenylcoumarin), 6,7-dihydroxy-4-methylcoumarin and imperatorin (8-isopentenyloxypsoralen) had strong activity among the coumarins examined. HL-60 cells treated with these coumarins differentiated into mature monocyte/macrophage. The structure-activity relationship established from the results revealed that a 6,7-dihydroxy moiety had an important role in the induction of differentiation of HL-60 (Kawaii et al., 2000).

Tumour cell-specific cytotoxicity was detected in 6,7-dihydroxy-substituted coumarins only, esculetin can be further enhanced by certain substitutions at 3- and/or 4-position(s) of the molecule. Esculetin and its derivatives with tumour-specific cytotoxicity induce internucleosomal DNA fragmentation in HL-60 cells (Kawase et al., 2003).

The mechanism of induction of apoptosis in HL-60 promyelocytic cells caused by scopoletin was investigated in a further study. Scopoletin induced apoptosis as confirmed by a characteristic ladder pattern of discontinuous DNA fragments in a dose-

dependent manner. It induced nuclear factor-kappa B activation, which, in turn, caused activation of caspase-3, degradation of poly(ADP-ribose) polymerase and finally DNA fragmentation, and eventually led to apoptotic cell death in HL-60 cells (Kim et al., 2005).

In a study twenty-one coumarins were examined for their antiproliferative activity towards several cancer cell lines. These authers also established that the 6,7-dihydroxy moiety had an important role for their antiproliferative activity in other cell lines (Kawaii et al., 2001).

Esculetin dose dependently inhibited the enhanced proliferation of cultured rabbit vascular smooth muscle cells stimulated by 5% foetal calf serum. The two adjacent phenolic hydroxyl groups at the C-6 and C-7 positions in the coumarin skeleton were necessary for the potent antiproliferative effect. The antiproliferative effects of other lipoxygenase inhibitors were comparable to the effect of esculetin. However, esculetin exhibited the greatest maximal suppression. The enhanced release of 12-hydroxyeicosatetraenoic acid, prostaglandin E2 and 6-keto-prostaglandin F1alpha in the culture medium of smooth muscle cells stimulated by 5% foetal calf serum were significantly reduced by esculetin. Furthermore, the foetal calf serum-stimulated protein tyrosine kinase activity was reduced by esculetin in a dose-dependent manner. In contrast, the protein kinase C activity stimulated by phorbol-12-myristate-13-acetate was not affected by esculetin. These results suggest that the antiproliferative effect of esculetin on vascular smooth muscle cells may be partly mediated through inhibition of protein tyrosine kinase and lipoxygenase (Huang et al., 1993).

The effects of a concomitant administration of esculetin and taxol were investigated in human hepatoma HepG2 cells. Firstly, esculetin alone could exert an antiproliferation effect together with an inhibitory effect on the activation of extracelluar signal-regulated kinase (ERK) and p38 mitogen-activated protein kinase. As compared to the treatment with taxol only, a co-administration with esculetin and taxol could result in a further enhancement of apoptosis, increase the expression of Bax (proapoptotic protein) and the cytosolic release of cytochrome C and enhance the expression of Fas and Fas ligand (apoptosis regulators) while the activation of caspase-8 and caspase-3 (components of apoptosis signalling cascade) was also increased. Finally, the ERK

cascade was proven to be involved in the enhancement of esculetin on the taxol-induced apoptosis (Kuo et al., 2006).

Scopoletin exerted a dual action on tumoral lymphocytes exhibiting both a cytostatic and a cytotoxic effect. These effects were associated with the induction of apoptosis. In contrast, scopoletin induced cell proliferation in normal T lymphocytes; this stimulatory action was found to be due to the interaction with Protein kinase C (PKC) (Manuele et al., 2006).

The dihydroxylated coumarins, 7,8-dihydroxy-coumarin and esculetin have a potential therapeutic role in the treatment of hematological malignancies. They induced DNA fragmentation as well as characteristic morphological changes of programmed cell death in leukemic cells. Results showed that the presence of two adjacent phenolic hydroxyl groups was the most important factor in terms of the structure-activity relationship. Radical semiquinone was detected in leukemic cells by electron spin resonance spectroscopy. The reactive oxygen species generation plays a critical role in dihydroxycoumarin-induced apoptosis in leukemic cells (Riveiro et al., 2008).

The mechanism of esculetin-induced anti-proliferative action and apoptosis was studied in human leukemia U937 cells. It was found that esculetin inhibits cell viability by inducing apoptosis, as evidenced by the formation of apoptotic bodies, DNA fragmentation, and the accumulation of cells in the sub-G1 phase. Esculetin-induced apoptosis was correlated with mitochondrial dysfunction, as well as the proteolytic activation of caspases. Furthermore, esculetin selectively increased the phosphorylation of extracellular signal-regulated kinase and c-Jun N-terminal kinase (components of apoptosis signalling cascade) (Park et al., 2008).

In general these studies emphasise the importance of adjacent hydroxyl groups on the coumarin skeleton. Many of the complex biological outcomes seen may depend on the known lipoxygenase inhibitory activity of esculetin and the complex signalling pathways leading to apoptosis. Relatively little work has explored the direct antioxidant effect of coumarins on oxidative damage to DNA. Oxidative stress results in DNA damage. 8-Oxo-7,8-dihydroguanine is the major oxidative DNA adduct. Esculetin and esculin inhibit 1,2-dimethylhhydrazine-induced formation of 8-Oxo-7,8-dihydroguanine so they have inhibitory effect on rat colon carcinogenesis (Kaneko et al., 2007).

Hypotensive and spasmolytic

Scopoletin has been identified as an active principle in the traditional herbal infusion of the fruit of *Tetrapleura tetraptera* TAUB used in the ethnopharmacology of West Africa as a potent hypotensive and non-specific spasmolytic agent.

Scopoletin isolated from *Atemisia afra* at a dose of 1.0-2.5 mg, induced a dose-dependent decrease in inotropic activity in rabbits and an appreciable decrease in chronotropic effects, especially at higher dose levels (Guantai and Addae-Mensah, 1999).

The hypotensive action of scopoletin is neither exerted through histamine H_1 receptor stimulation nor through α-adrenoceptor blockage and not mediated via the cholinergic mechanism but due to a non-specific spasmolytic action on the smooth muscle since it inhibited the spasmogenic effects of various agonists (Ojewole and Adesina, 1983).

Scopoletin (26-520 μM) inhibited to approximately the same extent the contractions induced by a variety of substances, including phenylephrine, potassium chloride, serotonin and $PGF_{2\alpha}$ in isolated rat aortic rings. This non-specific spasmolytic action of scopoletin can be attributed, at least in part, to its ability to inhibit intracellular calcium mobilization from noradrenaline-sensitive stores (Oliveira et al., 2001).

Other activities

Tyrosinase plays a critical regulatory role in melanin biosynthesis. Esculetin exhibited inhibition of mushroom tyrosinase activity against the oxidation of 3-(3,4-dihydroxyphenyl)-alanine by mushroom tyrosinase. The structure-activity relationships suggested that hydroxyl groups at the C_6 and C_7 positions of the coumarin skeleton played an important role in this inhibition (Masamoto et al., 2003). Esculetin 5 μM significantly suppressed melanin production in murine B16 melanoma cells without affecting cell growth. Furthermore, the number of 3,4-dihydroxyphenylalanine (DOPA)-positive melanocytes in the split-epidermal sheets treated with 0.05% or 0.1% esculetin was significantly lower than that in the control (Masamoto et al., 2004).

Scopoletin and its glucoside scopolin emerged as potential acetylcholinesterase inhibitors. They were isolated from *Scopolia carniolica* Jaqc and tested in an enzyme assay. They showed moderate, but significant, dose-dependent and long-lasting inhibitory activities. In *in vivo* experiments both increased the extracellular acetylcholine (Ach) concentration in rat brain to about 170% and 300% respectively compared to basal release. At the same concentration, the positive control galanthamine increased the ACh concentration to about the same level as scopoletin. These were the first *in vivo* results indicating an effect of coumarins on brain Ach (Rollinger et al., 2004).

Scopoletin inhibited hepatic lipid peroxidation and increased the activity of antioxidants, superoxide dismutase and catalase. Compared with the standard antithyroid drug, propylthiouracil, scopoletin exhibited a superior therapeutic activity, since unlike propylthiouracil, it also inhibited hepatic lipid peroxidation. These findings indicated that scopoletin had the potential to inhibit thyroid function and hyperglycemia without hepatotoxicity (Panda and Kar, 2006).

Scopoletin and its semisynthetic derivatives, acetylscopoletin and benzoylscopoletin, exhibited significant and dose-related antinociceptive effects against acetic acid-induced visceral pain. This might be useful in the development of new analgesic drugs for the management of visceral and inflammatory pain (Meotti et al., 2006).

Hypertriglyceridemia is caused by the imbalance between hepatic triglyceride production and peripheral removal. Lipoprotein lipase (LPL) plays a central role in the removal of plasma triglyceride. Scopoletin significantly increased LPL activity in adipocytes. Scopoletin increased LPL activity by increasing the LPL mRNA. These results suggested the possible action of scopoletin as a facilitator of plasma triglyceride clearance (Yang et al., 2007).

The effects of plant polyphenols on the activity of glutathione reductase (GSH-RD) were studied. Potency varied in the following order: chalcones > tannic acid > flavonoids > coumarins > catechins. C-6 and C-7 hydroxylations in coumarins were important for their inhibition activities of GSH-RD. The inhibition of GSH-RD by esculin was reversible but not time dependent (Zhang et al., 1997).

Scopoletin from *A. dahurica* protects neuronal cells from the damage caused by oxygen and glucose deprivation. It reduced the propidium iodide uptake, which is an indication of impaired cell membrane integrity. In addition, it inhibited the loss of NeuN (neuronal-specific nuclear protein), which represents the viability of neuronal cells (Son et al., 2007).

The chondroprotective effect of esculetin was investigated using primary cultures of rabbit articular chondrocytes. Although esculetin did not directly inhibit collagenolytic activity in the culture media, it significantly suppressed the production of pro-matrix metalloproteinase-1/interstitial procollagenase and pro-matrix metalloproteinase-3/prostromelysin 1, accompanied by a decrease in the steady-state levels of their mRNAs. These results suggest that esculetin could be a therapeutically effective candidate for inhibition of cartilage destruction in osteoarthritis and rheumatoid arthritis (Watanabe et al., 1999).

Esculetin, was predicted to be potentially useful as an ingredient in cosmetics for protecting against photoaging. Esculetin has antioxidant activity and an inhibitory effect on the expression of matrix metalloproteinases induced by UV irradiation, in addition to showing scavenging activities for DPPH radicals and superoxide anions generated from xanthine oxidase (Lee et al., 2007).

Glycosyltransferases

Glucosylation is considered to be the final step in the biosynthesis of scopolin (Fig. 1.13). Glycosyltransferases (GTs) are found in all living organisms (eukaryote and prokaryote) where they catalyse the conjugation of a hydroxyl group (alcoholic or carboxylic) with sugars. They use nucleotide sugars e.g. UDP-glucose (uridine diphosphate glucose) in plants and in mammals as glucose donor and UDP-glucuronic acid in mammals as glucuronic acid donor; UDP is the (biological) leaving group. The term glycosyltransferases signifies a class of enzymes that catalyse the transfer of a monosaccharide from the sugar nucleotide to the aglycone (without sugar) to form a glycoside. Most are highly specific to the monosaccharide unit and often specific for a small range of aglycones. Thus glucosyltransferases catalyse the transfer of glucose (specifically) to the aglycone to form a glucoside (Fig. 1.13).

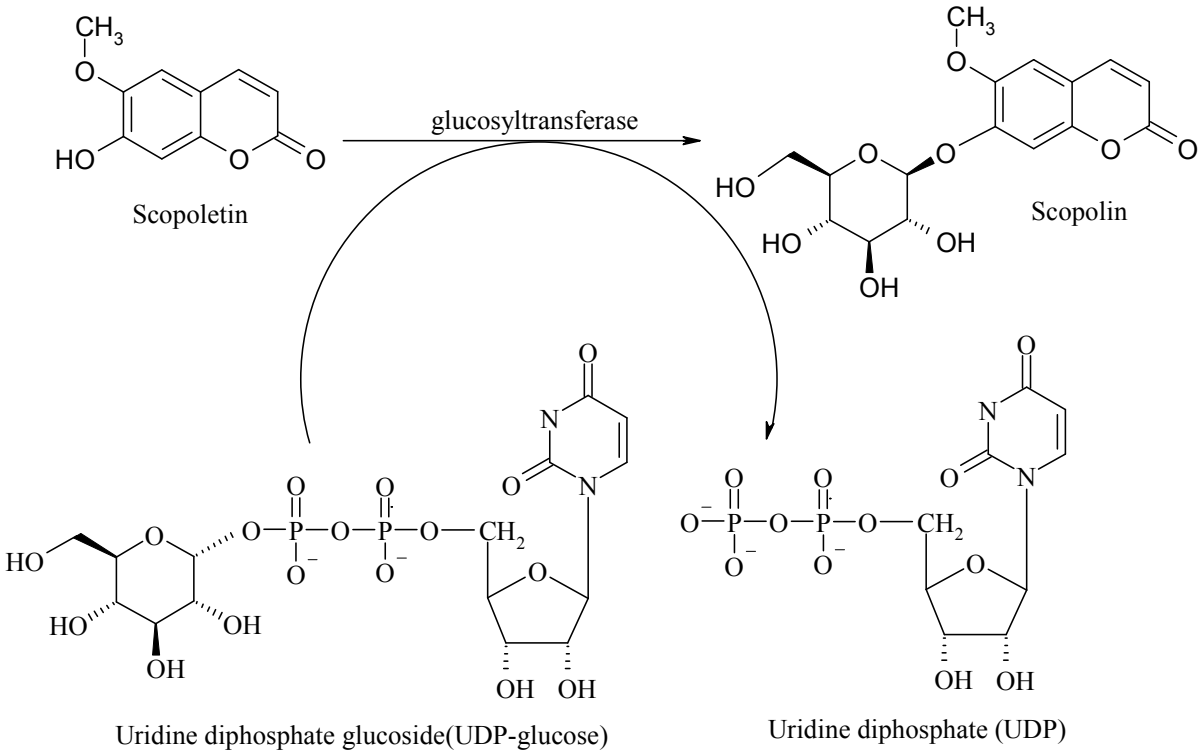


Fig. 1.13. Biosynthesis of scopolin from scopoletin.

In higher plants, secondary metabolites are often converted into their glycoconjugates, which then accumulate in the vacuole. The roles of glycosylation in plants include: solublisation of the compound in water; detoxification of harmful metabolites or environmental compounds such as herbicides; and the regulation of the action of functional compounds, such as plant hormones (Bowles et al., 2005). Also, glucosyltransferases regulate the activity of compounds that play important roles in plant defence responses against pathogens such as tobacco mosaic virus (Chong et al., 2002) and potato virus (Matros and Mock, 2004). Transgenic tobacco plants that over-expressed glucosyltransferase showed a 2-fold higher accumulation of scopoletin and scopolin after inoculation with mosaic virus compared to wild-type plants. Necrosis appeared sooner in transgenic plants and lesions developed faster, suggesting an accelerated hypersensitive response. Unexpectedly, the viral content in each lesion was not significantly different in transgenic compared to wild-type plants (Gachon et al., 2004).

Recently GTs have been divided into families according to sequence identity and the nature of the biochemical reaction; usually, within a given family, the various enzymes transfer the same sugar residue (http://afmb.cnrs-mrs.fr/CAZY/GT.html) (Thorsoe et al., 2005). The current number of GT families includes 91 families in the CAZy (denotes Carbohydrate-Active enzymes) data base http://www.cazy.org/fam/acc_GT.html). The recommended nomenclature has been updated (Mackenzie et al., 1997); for each gene, it is recommended that the root symbol *UGT* (*Ugt* for mouse and *Drosophila*), denoting "UDP glycosyltransferase", be followed by an Arabic number representing the family, a letter designating the subfamily, and an Arabic numeral denoting the individual gene within the family or subfamily, e.g. 'human *UGT*2B4' and 'mouse *Ugt*2b5'. As with the P450 nomenclature, families 1 – 50 are reserved for animals, 51 – 70 for yeast, 71 – 100 for plants and 101 – 200 for bacteria. For cDNA, mRNA, protein and enzyme activity, it is recommended to have the same nomenclature of its gene or it could be non italicized, all capitals even for mouse and *Drosophila*, for an example UGT71A1 is the cDNA of *UGT*71A1 in cassava.

Plant GTs have been reviewed recently (Vogt and Jones, 2000; Jones and Vogt, 2001; Lim and Bowles, 2004). Additionally, the UDP-glycosyltransferases of *Arabidopsis thaliana* has been analysed, and phylogenetic trees have been constructed (Li et al., 2001). Conserved domain structures for substrate recognition and nucleotide-sugar binding are expected to be located within the enzyme (Kapitonov and Yu, 1999).

The complete sequence of the *A. thaliana* genome enables definitive characterization of multigene families and analysis of their phylogenetic relationships so as to explore the evolution of substrate recognition and regioselectivity of glucosylation. Hydroxycoumarins (scopoletin and esculetin) have been used as model substrates for an analysis in which 90 sequences were expressed and the resulting enzymes were assayed and of which 48 were shown to recognize these compounds. The study revealed activity in 6 of the 14 phylogenetic groups of the multigene family, suggesting that basic features of substrate recognition are retained across substantial evolutionary periods (Lim et al., 2003).

Hydroxycoumarin 7-*O*-glucosyltransferase, which catalyses the formation of scopolin from scopoletin, was purified from a culture of tobacco cells. The purified

enzyme had a pH optimum of 7.5, an isoelectric point (pI) of 5.0, and a molecular mass of 49 kDa. The enzyme did not require metal cofactors for activity. The purified enzyme preferred hydroxycoumarins as substrates but also exhibited significant activity with flavonoids. A database search using the amino terminus amino acid sequence of this glucosyltransferase revealed strong homology to the amino acid sequences of other glucosyltransferases in plants (Taguchi et al., 2000).

All UDP glycosyltransferases contain a 44-amino acid region of highly conserved homology in their carboxyl-terminal halves. This region is thought to be involved in binding to the UDP moiety of the nucleotide sugar (Hundle et al., 1992; Mackenzie et al., 1997). Moreover, in a comparison of the amino acid sequences of seven species of glucosyltransferases with six species of galactosyltransferases, glutamine and histidine were found to be highly conserved as the last amino acid residue of a glycosyltransferase-specific conserved region (UGT) in glucosyltransferases and galactosyltransferases, respectively. Consequently, the sugar donor specificities of glycosyltransferases can be successfully altered by a single amino acid via a point mutation (Kubo et al., 2004). On the other hand *in vitro* studies have shown that a single gene product can glycosylate multiple substrates of diverse origins; multiple enzymes can also glycosylate the same substrate (Bowles et al., 2006). These features suggest that in a cellular context, substrate availability is a determining factor in enzyme function, and redundancy depends on gene regulation (Bowles et al., 2006).

In the biosynthesis of the cyanogenic glucoside dhurrin in *Sorghum bicolor*, the UDP-glucosyltransferase *UGT*85B1 catalyzes the conversion of *p*-hydroxymandelonitrile into dhurrin. A structural model of *UGT*85B1 has been built that enables predictions about amino acid residues important for catalysis and sugar donor specificity to be made. *p*-Hydroxymandelonitrile and UDP-glucose (Glc) were predicted to be positioned within hydrogen-bonding distance of a glutamic acid residue in position E410 facilitating sugar transfer. The acceptor was packed within van der Waals distance to histidine H23. Serine S391 and arginine R201 form hydrogen bonds to the pyrophosphate part of UDP-Glc and hence stabilize binding of the sugar donor (Thorsoe et al., 2005).

Aims of the work

The overall aim of this project was to conduct integrated studies at the phytochemical, biosynthetic and genetic levels into a single important crop species, cassava. This involved the phytochemical comparison of secondary metabolite production in fresh and deteriorated cassava roots in order to further understanding of this aspect of post-harvest physiological deterioration (PPD) that causes significant wastage and economic loss. Ultimately, the results from these studies may help to contribute towards the control of this problem.

Hydroxycoumarins accumulate in cassava roots during PPD and play important roles in plant development and defence; additionally, some are pharmacologically active, but key aspects of the biosynthesis of these secondary metabolites remain unresolved. Thus, this project exploits the accumulation of scopoletin and scopolin in cassava roots during PPD to test alternative pathways for the biosynthesis of these hydroxycoumarins using stable isotope labelling. Scopolin is the glucoside of scopoletin and is one of many important glycosides in cassava.

The first objective is a phytochemical study of secondary metabolites in healthy and deteriorated cassava roots by isolation of active constituents by different chromatographic methods and identification of the isolated compounds by studying their physical and chemical properties on the basis of chromatographic separation and using a variety of modern spectroscopic analytical techniques (mainly NMR and HR MS). Then the biosynthetic pathway of scopoletin and its glucoside scopolin will be studied in cassava root during PPD. Using different, designed and chemically synthesised isotopically labelled intermediates along the biosynthetic pathway, the *E*-*Z*-isomerisation step, o-hydroxylation and lactonisation will be investigated.

Another major objective is the identification, isolation and characterisation of glucosyltransferases (GTs) in a cDNA library obtained from deteriorated cassava roots, including those GT enzymes involved in the biosynthesis of scopolin from scopoletin and comparing their expression in fresh and deteriorated cassava roots. This will provide clear evidence that these genes are expressed in the cassava storage root during PPD. Using the bioinformatics technique of Blast searching the NCBI database, it will be shown if these GTs are novel or have been previously isolated.

Chapter 2

Identification of secondary metabolites in fresh and deteriorated cassava roots

2.1. Introduction

Cassava (*Manihot esculenta*) tuberous roots are a staple food for more than five hundred million people in tropical countries due to their high starch content and ability to withstand unfavourable conditions. The crop suffers from some problems that limit its utilization such as the roots have a short shelf life due to post-harvest physiological deterioration (PPD). Within 2-3 days after harvesting the root shows blue to black (Fig. 2.1) vascular streaking and it becomes unpalatable and therefore unmarketable, this PPD significantly affects the crop's economic value.

Studies show that within 15 min of the root being injured an oxidative burst occurs, which is followed by changes in the expression of genes and the accumulation of secondary metabolites (Sakai and Nakagawa, 1988; Reilly et al., 2004) e.g. the antioxidant hydroxycoumarins scopoletin and esculetin, and their respective glucosides scopolin and esculin (Buschmann et al., 2000b). Scopoletin may give rise to the blue/black discolouration by oxidation and polymerisation.

In order to more fully understand these phenomena, and with a view to overcoming this problem, the ethanolic extracts of fresh and deteriorated cassava roots were analyzed by chromatographic methods to isolate and then identify the constituents by a variety of spectrometric techniques.

Fig. 2.1. Fresh and deteriorated cassava roots. Deteriorated cassava roots show blue florescence under UV due to the accumulation of hydroxycoumarins.

2.2. Experimental

Plant material

Root tubers of different cultivars (MCOL 22, MNGA 19, MNGA 2) were harvested from cassava plants growing in the tropical glass house at the University of Bath under the following condition: temperature (22-28 °C), relative humidity (R.H.) (40-80%) and a light period of 14 hours per day. The cultivar identification code is the CIAT identification code relating to country of origin, for example MCOL 22 denotes *Manihot* Colombia accession number 22 while MNGA 19 denotes *Manihot* Nigeria accession number 19. The cultivar MCOL 22 shows high susceptibility to PPD, MNGA 2 shows medium susceptibility while MNGA 19 is not determined, according to CIAT.

General methods

All solvents used were supplied by Fisher, all in GPR grade, except where stated. Chemicals were obtained routinely from Sigma-Aldrich Chemical Co. Ltd, UK.

Thin Layer Chromatography (TLC)

Analytical TLC was performed on commercial aluminium-backed plates pre-coated with silica GF_{254} 60 in 0.25mm thickness, purchased from Merck. The mobile phases predominantly utilized were (solvents ratios are v/v/v unless otherwise stated): $CHCl_3$: EtOAc: MeOH 2: 2: 1; EtOAc: hexane: formic acid 1: 1: 0.1.

Diluted samples were spotted on to the plates and developed for varied time scales. The developed plates were dried and observed in visible light and under UV light at 254 nm and 365 nm. TLC plates were then sprayed with 1% (w/v) vanillin in conc. sulfuric acid and heated at 100 °C and observed in visible light.

Column Chromatography

Column chromatography on a preparative scale was carried out on silica gel (purchased from Merck, 40-63 μm, 230 – 400 mesh ASTM, pH 6.5 – 7.5 for a 10 % suspension). The column was packed in an open-glass column with an appropriate solvent of relatively lower polarity compared to the solvent system that could achieve a

satisfactory resolution on TLC. If samples were well dissolved in packing solvent, then they were applied to the column in concentrated solution in minimum volume. Samples with poor solubility in packing solvent were applied to the column in the solid phase dispersed on an appropriate amount of Celite.

The progress of the elution was followed by collecting fractions from the eluate, concentrating and monitoring by analytical TLC, HPLC and NMR. Appropriate combination of fractions was determined by their TLC behaviour.

High Performance Liquid Chromatography (HPLC)

The HPLC instrument was a solvent delivery system equipped with Jasco PU-980 pump and monitored at 360 nm by a Jasco UV-975 detector. The traces were recorded on a Goerz Metrawatt Servogor 120 recorder**.** All columns were commercially pre-packed reversed phase HPLC columns, purchased from Phenomenex Inc.:
Semi- Preparative Column: Phenomenex Gemini 10 µ C18 110A 250 x 10 mm;
Guard Column: Phenomenex Gemini 5 µ C18 10 ×10 mm.
Detection was by UV absorbance at 360 nm.

Acetonitrile (ACN) used was of HPLC grade supplied by Fisher. Solvents were filtered through Whatman Nylon membrane prior to use. Water was filtered by Milli-Q Plus PF. An appropriate percentage of organic solvent was mixed with water. The aqueous mobile system was shaken vigorously and degassed on a Decon ultrasonicater for 30 minutes before use. Samples were dissolved in pure organic solvent and filtered through 17mm SYR Filter Nylon 0.45µm before injection using a 100 µl loop. All the HPLC experiments were performed using isocratic elution.

Ultraviolet Spectroscopy (UV)

Ultraviolet spectra were measured, unless stated, in solution in methanol using a Perkin-Elmer UV/VIS spectrophotometer Lambda EZ201.

Mass Spectrometry (MS)

High Resolution Electrospray Ionisation Mass Spectrometry (HR ESI MS) was carried out on a Bruker micrOTOF mass spectrometer in the Department of Pharmacy

and Pharmacology, University of Bath. HR ESI MS and High Resolution Fast Atom bombardment (HR FAB MS) were carried out on a Micromass Quattro II in EPSRC National Mass Spectrometry Service Centre, University of Wales, Swansea.

Nuclear Magnetic Resonance Spectroscopy (NMR)

NMR spectra were obtained on either JEOL GX 270 (operating at 270 MHz for ^{1}H and 67.8 MHz for ^{13}C) or Varian Mercury 400 MHz (operating at 400 MHz for ^{1}H and 100 MHz for ^{13}C) spectrometers in $CDCl_3$, CD_3OD or pyridine-d_5 and all chemical shifts are reported in parts per million relative to internal tetramethylsilane. COSY, DEPT, HMQC, HMBC were all recorded at normal frequency. Coupling constants (*J* values, line separations) are absolute values reported in Hz and the multiplicities are abbreviated: s (singlet), d (doublet), t (triplet), q (quartet), m (multiplet) and br (broad).

General extraction method

Cassava roots were peeled then cut into approximately 1 cm^3 cubes and divided into groups. One group was crushed and extracted with EtOH (fresh cassava extract) and other groups allowed to deteriorate for up to 6 d under controlled conditions (20° C, 80-90% R.H.) until crushing and extraction. EtOH was used because it is a water miscible solvent with high polarity and has previously been reported to effectively extract hydroxycoumarins. Daily from day 3, approx. 270 g of the sliced roots were crushed and macerated (EtOH 200 ml). The ethanolic extracts were filtered and evaporated under reduced pressure at 35-40 °C. The EtOH extracts of these 5 samples were analysed by TLC ($CHCl_3$: EtOAc: MeOH 2: 2: 1, visualised by UV at 365 nm) using cholesterol and scopoletin as markers enabled appropriate column chromatography conditions to be selected to achieve optimized separation of target compounds. Also several HPLC solvent mixtures were selected to optimize the separation of target compounds and to afford pure samples.

Investigation of active constituents in cassava root after deterioration

Using the General Extraction Method, cassava roots (1.3 kg, various cultivars) were peeled (1.1 kg). Four samples of ethanolic extracts of deteriorated cassava roots of

day three, day four, day five and day six after harvesting cassava roots were prepared. EtOH extracts of these 4 samples were analysed by TLC ($CHCl_3$: EtOAc: MeOH 2: 2: 1, visualised by UV at 365 nm). These four extracts were combined. The combined extract was fractionated between water (see below) and $CHCl_3$.

The chloroformic fraction

The combined $CHCl_3$ layers were concentrated to yield a pale yellow viscous oil (240 mg) which was then further purified by solid-phase extraction. Several times 1 g cartridges (C_{18}) (Varian) typically were washed (MeOH, 5 ml) and conditioned (50% aq. MeOH, 5 ml) and 8 ml of the extract (1 mg/ml in MeOH with sufficient water added to the solution until it became opalescent) was loaded and eluted with 50% aq. MeOH. The eluant was concentrated under reduced pressure which afforded scopoletin, identical by TLC to a commercial sample. The residue was dissolved in MeOH and purified by semi-preparative HPLC.

The HPLC peak at 24.7 min was collected from injections (100 µl), combined, concentrated to yield compound (1) (4 mg). This was identified by high field NMR spectroscopy and High Resolution Electrospray Ionisation Mass Spectrometry (HR ESI MS) (Hirata et al., 2000), as the coumarin, scopoletin. ^{1}H NMR: δ 3.90 (3H, s, 6-OCH_3), 6.19 (1H, d, J = 9, H-3), 6.75 (1H, s, H-8), 7.09 (1H, s, H-5), 7.84 (1H, d, J = 9, H-4). HR MS: $C_{10}H_9O_4$ requires 193.0495, HR MS found *m/z* 193.0489 $[M + H]^+$, $C_{10}H_8O_4Na$ requires 215.0315, found 215.0309 $[M + Na]^+$.

The aqueous fraction

The concentrated aq. fraction (7.6 g) was separated using the above HPLC conditions. HPLC peaks were observed at 4.5, 6.0, 7.1 and 24.7 min.

The HPLC peak at 4.5 min was collected from 100 injections (100 µl), combined and concentrated to yield compound (2) (40 mg). ^{1}H NMR (CD_3OD): δ 1.66 (3H, s, H-1), 1.67 (3H, s, H-3), 3.18-3.76 (5H overlapping, H-2',3',4',5',6'a), 3.85 (1H, dd, J = 11.2, 1.7 Hz, H-6'b), 4.63 (1H, d, J = 7.8 Hz, H-1'). ^{13}C NMR (CD_3OD): δ 28.45 (C-1), 29.20 (C-3), 63.23 (C-6'), 72.00 (C-4'), 73.44 (C-2), 75.45 (C-2'), 78.55 (C-5'), 78.66 (C-3'),

101.89 (C-1'), 122.90 (CN). From these spectral data, compound 2 was concluded to be linamarin.

The HPLC peak at 6.0 min was collected from 100 injections (100 µl), combined and concentrated to yield compound (3) (1 mg). It was identified as esculin by comparison of UV and HPLC retention time with an authentic sample.

The HPLC peak at 7.1 min was collected from 100 injections (100 µl), combined and concentrated to yield compound (4) (7 mg). It was identified by high field NMR spectroscopy and HR MS. ^{1}H NMR (CD_3OD): δ 3.41-3.55 (4H overlapping, H-2',3',4',5'), 3.71 (1H, br s, H-6'a), 3.84 (1H, br s, H-6'b), 3.91 (3H, s, 6-OCH_3), 5.09 (1H, d, J = 8, H-1'), 6.31 (1H, d, J = 9, H-3), 7.18 (1H, s, H-8), 7.22 (1H, s, H-5), 7.91 (1H, d, J = 9, H-4). ^{13}C NMR (CD_3OD): δ 55.8 (6-OCH_3), 61.2 (C-6'), 70.0 (C-4'), 73.5 (C-2'), 76.6 (C-5'), 77.2 (C-3'), 100.8 (C-1'), 104.0 (C-8), 109.5 (C-5), 113.3 (C-3, C-4a), 144.5 (C-4), 147.0 (C-6), 149.5 (C-8a), 150.5 (C-7), 162.4 (C-2). HR MS: Naturally occurring scopolin $C_{16}H_{19}O_9$ requires 355.1024, HR MS found *m/z* 355.1033 $[M + H]^+$, $C_{16}H_{18}O_9Na$ requires 377.0843, found 377.0860 $[M + Na]^+$.

The HPLC peak at 11.5 min was collected from 100 injections (100 µl), combined and concentrated to yield compound (5) (1 mg). It was identified as esculetin by comparison of UV and HPLC retention time with an authentic sample.

The HPLC peak at 24.7 min was collected from 100 injections (100 µl), combined and concentrated to yield (2 mg). It was identified as scopoletin by comparison of UV and HPLC retention time with an authentic sample and with the scopoletin compound (1) identified from $CHCl_3$ fraction above.

Investigation of constituents in healthy cassava roots

Using the general extraction method, cassava roots (5 kg, various cultivars) were peeled (4.3 kg) and extracted immediately with ethanol which was then concentrated (109 g). This was extracted with hot water and water-insoluble material was collected by centrifugation. The water insoluble fraction (8.7 g) was chromatographed over silica gel (120 g) and was eluted with a gradient of 100% EtOAc, 10%, 25% and 50% MeOH in 100% EtOAc (1.5 L of each) to give 24 fractions which were monitored by TLC (Merck, Kieselgel 60 F_{254}) using $CHCl_3$-EtOAc-MeOH (2:2:1) and in EtOAc and detected by 1 %

vanillin in sulphuric acid reagent, and by 270 MHz ^{1}H NMR. Fraction 2 was analysed by UV and compound (6) was identified as β-carotene.

Fractions 8-10 (600 mg) were combined and chromatographed on silica gel (60 g) and eluted with EtOAc-CH_2Cl_2-MeOH (1l, 2:2:0.5) to give 31 fractions, fr 10 was compound (7). ^{1}H NMR (pyridine-d_5, 400 MHz): δ 0.82 – 0.86 (6H, m, 2 x methyl), 1.23–1.33 (m, chain methylenes), 1.60 (4H, m, 2 x β-methylene), 2.05-2.09 (8H, m, 4 x allylic methylene), 2.30–2.33 (4H, m, 2 x α-methylene), 2.89 (2H, t, *J* = 6, methylene), 4.05 (2H, t, *J* = 6, H-3a, H-5'), 4.14 (1H, dd, *J* = 3, 9, H-3'), 4.35 (1H, dd, *J* = 5, 11, H-3b), 4.43 (3H, d, *J* = 9, H-2', 6'), 4.50 (1H, dd, *J* = 7, 12, H-1b), 4.55 (1H, d, *J* = 3, H-4'), 4.67 (1H, dd, *J* = 5, 12, H-1a), 4.81 (1H, d, *J* = 8, H-1'), 5.43–5.52 (m, olefinic), 5.64–5.66 (1H, m, H-2). ^{13}C NMR (pyridine-d_5, 100 MHz): δ 14.2 (methyl), 14.3 (methyl), 22.8 (methylene), 22.9 (methylene), 25.2 (β-methylene), 26.0 (methylene), 27.4 (allylic methylene), 27.5 (allylic methylene), 29.3-30.0 (methylene), 31.6 (methylene), 32.1 (methylene), 34.1 (α-methylene), 34.4 (α-methylene), 62.2 (C-6'), 63.2 (C-1), 68.0 (C-3), 70.0 (C-4'), 70.9 (C-2), 72.2 (C-2'), 75.2 (C-3'), 77.1 (C-5'), 105.7 (C-1'), 128.3 (olefinic), 128.4 (olefinic), 130.1 (olefinic), 130.2 (olefinic), 130.3 (olefinic), 130.4 (olefinic), 173.1 (olefinic), 173.3 (olefinic). HR FAB MS: *m/z* of $[M+Na]^+$ $C_{45}H_{80}O_{10}Na$ requires 803.6, found 803.6 and $C_{45}H_{82}O_{10}Na$ requires 805.6, found 805.6.

Fractions 12-14 (95 mg) were combined and chromatographed on silica gel (20 g) and eluted with 500 ml of EtOAc-CH_2Cl_2-MeOH (2:2:0.5) to give compound (8). ^{1}H NMR (pyridine-d_5, 400 MHz): δ 0.63 (3H, s, Me-18), 0.83-0.90 (9H, m 3 x Me-26, 27, 29), 0.91 (3H, s, Me-19), 0.96 (3H, d, *J* = 7, Me-21), 0.99-2.73 (36H, m, methylene protons), 3.91-3.99 (1H, m, H-3), 4.05 (1H, t, *J* = 8, H-2'), 4.28 (2H, m, H-3', 4'), 4.39 (1H, dd, *J* = 12, 5, H-6'b), 4.55 (1H, dd, *J* = 12, 2, H-6'a), 5.04 (1H, d, *J* = 8, H-1'), 5.32-5.33 (1H, m, H-6). ^{13}C NMR (pyridine-d_5, 100 MHz): δ 12.0 (C-18), 12.1 (C-29), 19.0 (C-21), 19.2 (C-19), 19.4 (C-27), 20.0 (C-26), 21.3 (C-11), 23.4 (C-28), 24.5 (C-15), 26.4 (C-23), 28.5 (C-16), 29.4 (C-25), 30.2 (C-2), 32.0 (C-8), 32.2 (C-7), 34.2 (C-22), 36.4 (C-20), 36.9 (C-10), 37.5 (C-1), 39.3 (C-4), 39.9 (C-12), 42.5 (C-13), 46.0 (C-24), 50.3 (C-9), 56.2 (C-17), 56.8 (C-14), 62.8 (C-6'), 71.7 (C-4'), 75.3 (C-2'), 78.1 (C-3), 78.5 (C-5'), 78.6 (C-3'), 102.6 (C-1'), 121.9 (C-6), 140.9 (C-5). HR ESI MS: *m/z* of $[M+NH_4]^+$ $C_{35}H_{64}O_6N_1$ requires 594.4728, found 594.4733.

2.3. Results and discussion

Comparison between fresh and deteriorated ethanolic cassava root extracts

The ethanolic extracts of fresh and deteriorated cassava roots were analysed by TLC ($CHCl_3$: EtOAc: MeOH 2: 2: 1) (Fig. 2.2) and there was dramatic accumulation of hydroxycoumarins which can be visualised as blue coloured spots under UV at 365 nm. No clear difference was observed between days 3 and 4 deteriorated ethanolic extracts, where it is known that the highest amount of scopoletin and scopolin accumulation occurs (Buschmann et al., 2000b). Scopoletin and its glucoside, scopolin were accumulated to a greater extent than esculetin and its glucoside, esculin. This TLC confirmed that the accumulation of hydroxycoumarin could be related to PPD and that interfering with their biosynthesis could be a way for increasing the shelf life of harvested root.

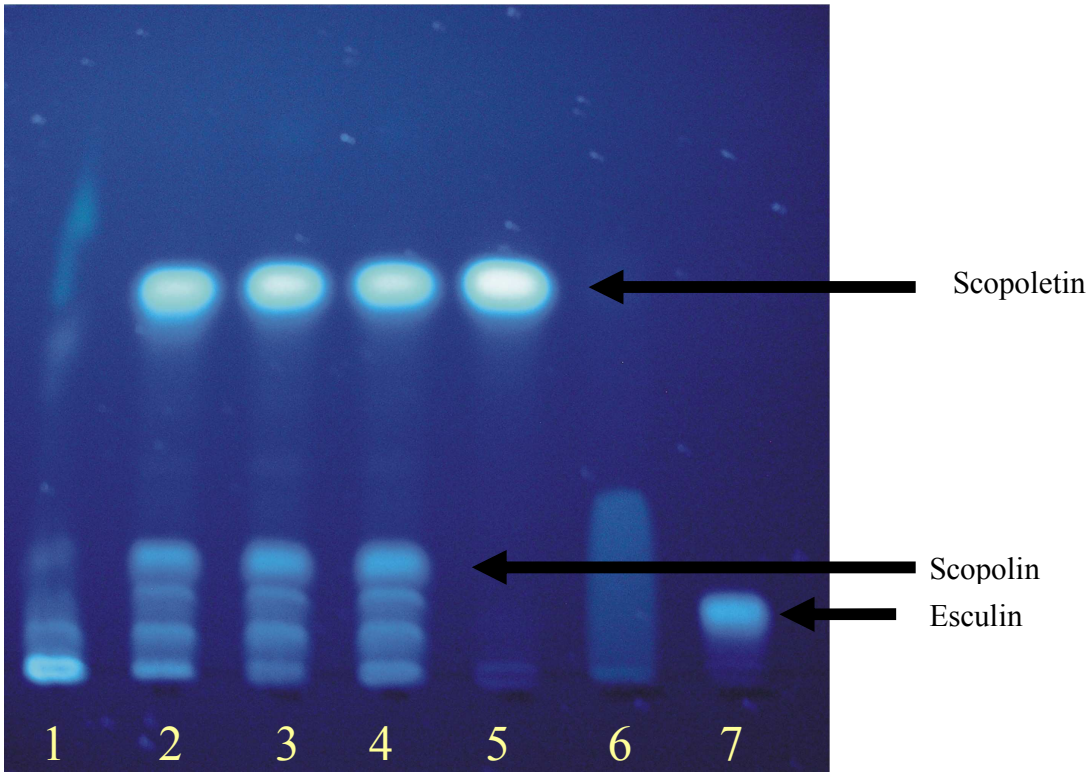


Fig. 2.2. TLC of cassava EtOH extracts, lane 1 is fresh, lanes 2-4 are day 3, 4, and 5 deteriorated extracts, while lanes 5, 6 and 7 are authentic scopoletin, esculetin (tailing) and esculin respectively, TLC solvent system is $CHCl_3$-EtOAc-MeOH (2:2:1 v/v/v), detected under UV at 360 nm. Scopolin identified by isolation and spectroscopic analysis.

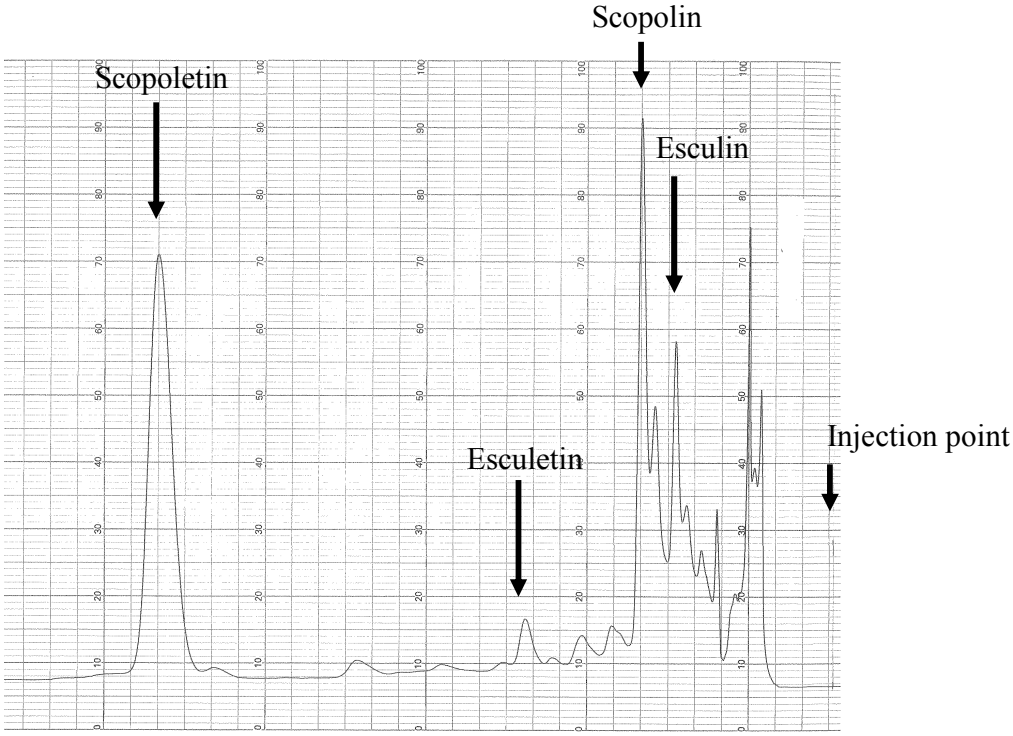


Fig. 2.3. HPLC trace of the ethanolic extract of deteriorated cassava roots, the peaks were identified by comparing their retention times with authentic standards.

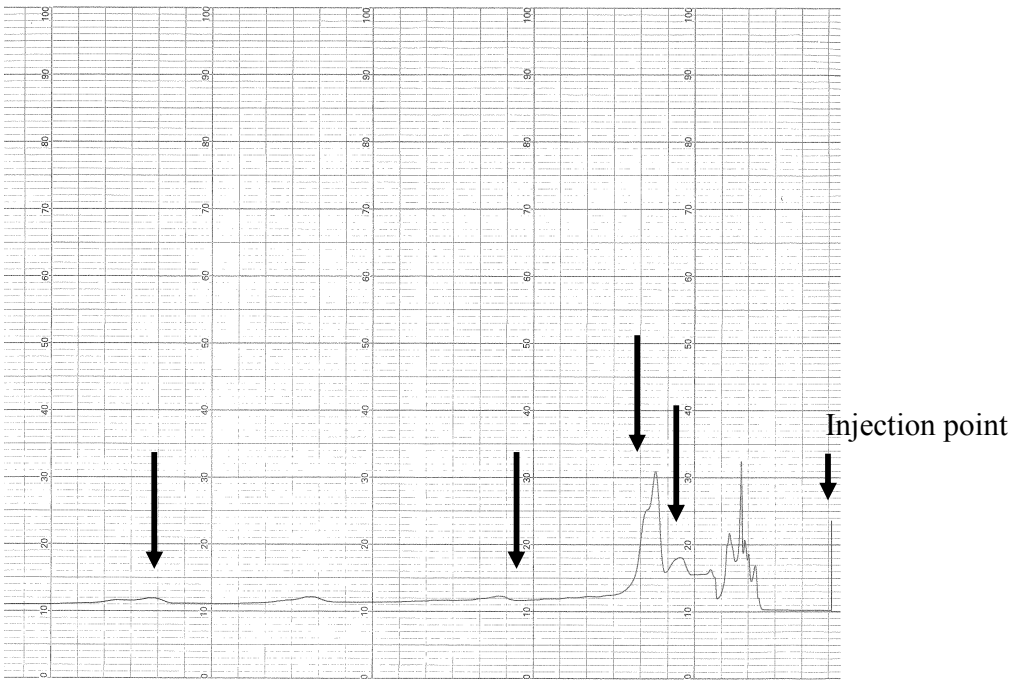


Fig. 2.4. HPLC trace of the ethanolic extract of fresh cassava roots, the peaks were identified by comparing their retention times with authentic standards.

The ethanolic extracts of fresh and deteriorated cassava roots were also analysed by HPLC (Fig. 2.3 and 2.4) using 16% acetonitrile in water containing 0.1% formic acid, flow rate 4 ml/min. HPLC traces showing the high accumulation of hydroxycoumarins in the deteriorated extract in particular scopoletin and scopolin in comparison with the fresh extract which shows a very small amount of these hydroxycoumarins

Investigation of constituents in cassava root after deterioration

Compound (1) (4 mg, isolated from the chloroformic fraction of the ethanolic extract of deteriorated cassava root) was identified as scopoletin by high field NMR spectroscopy, HR MS and comparison with published data (Fig. 2.5-2.7) (Hirata et al., 2000). ^{1}H NMR showed four peaks in the aromatic region; two of them are doublets of J = 9 Hz the more deshielded signal at δ 7.84 is H-4 due the deshielding effect of the benzene ring while the less deshielded at 6.19 is H-3. The other two signals are singlets at 6.75 and 7.09 are H-8 and H-5 in addition to the OCH_3 singlet peak at 3.90. Identification was also confirmed by comparison of the HPLC retention time at 24.7 min with authentic samples of scopoletin and isoscopoletin (Fig.2.8).

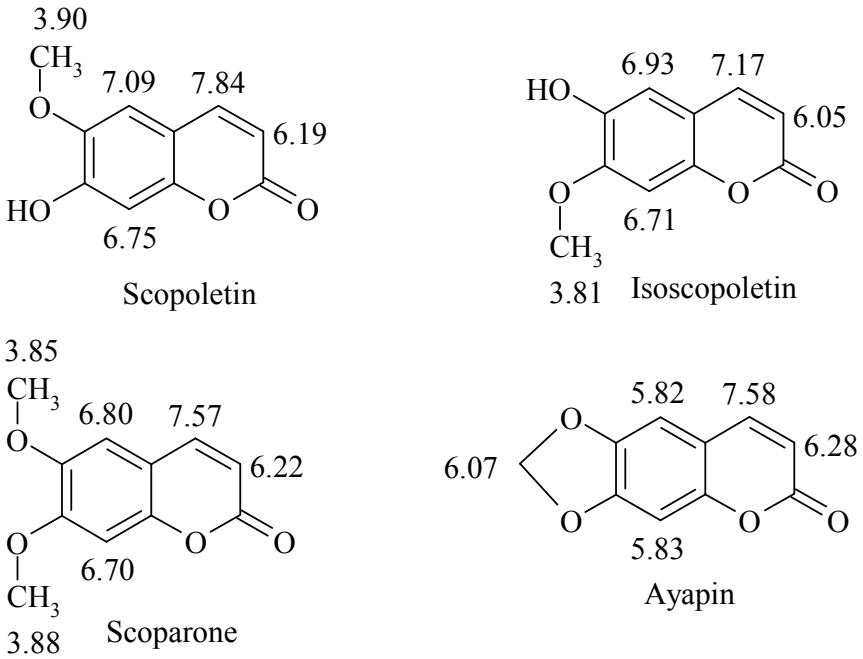


Fig. 2.5. Comparison of ^{1}H NMR data for scopoletin (CD_3OD) with reported ^{1}H NMR data ($CDCl_3$) of isoscopoletin, scoparone (6,7-dimethoxycoumarin) from *Artemesia tridentata* spp. *vaseyana* (Brown et al., 1975) and ayapin (the corresponding methylenedioxy) from *Helianthus annuus* (Gutierrez et al., 1995; Maes et al., 2005).

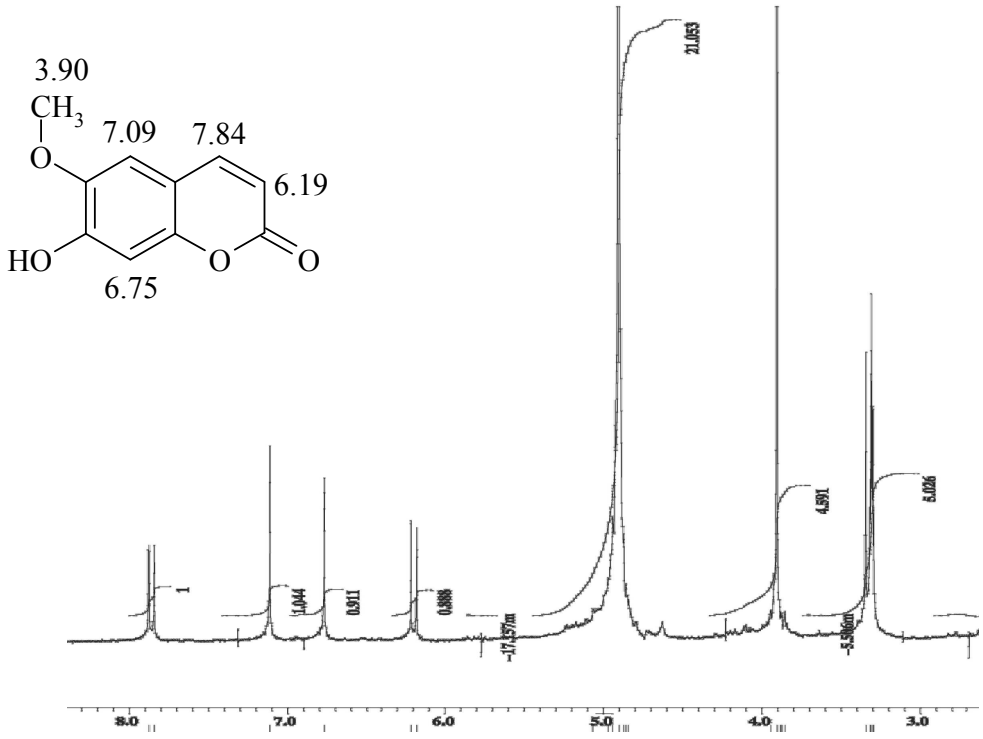


Fig. 2.6. ^{1}H NMR spectrum (CD_3OD, 400 MHz) of scopoletin (1).

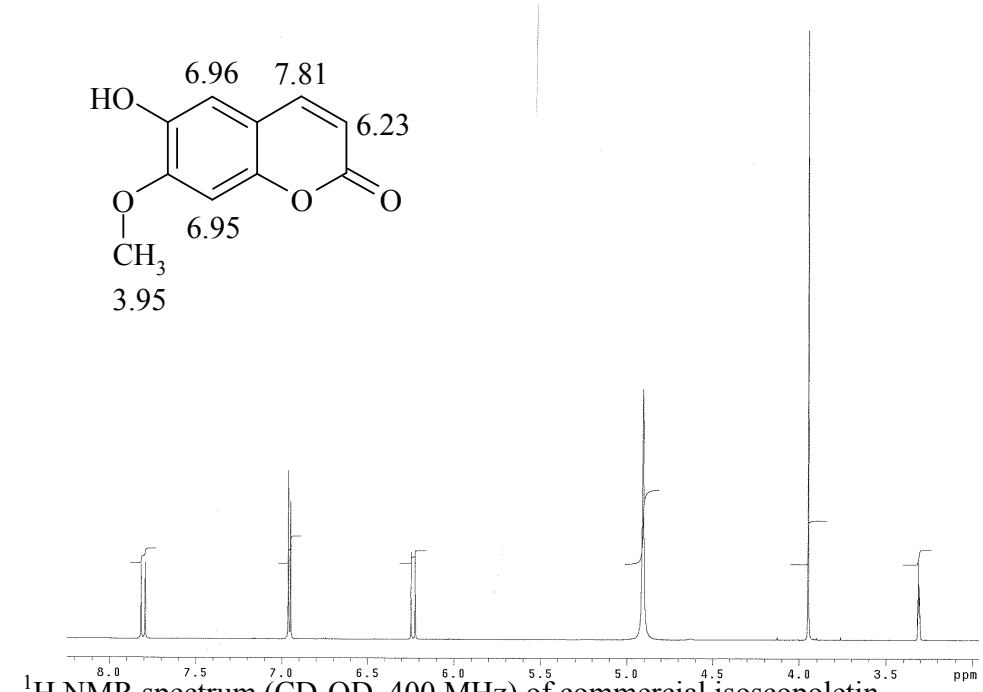


Fig. 2.7. ^{1}H NMR spectrum (CD_3OD, 400 MHz) of commercial isoscopoletin.

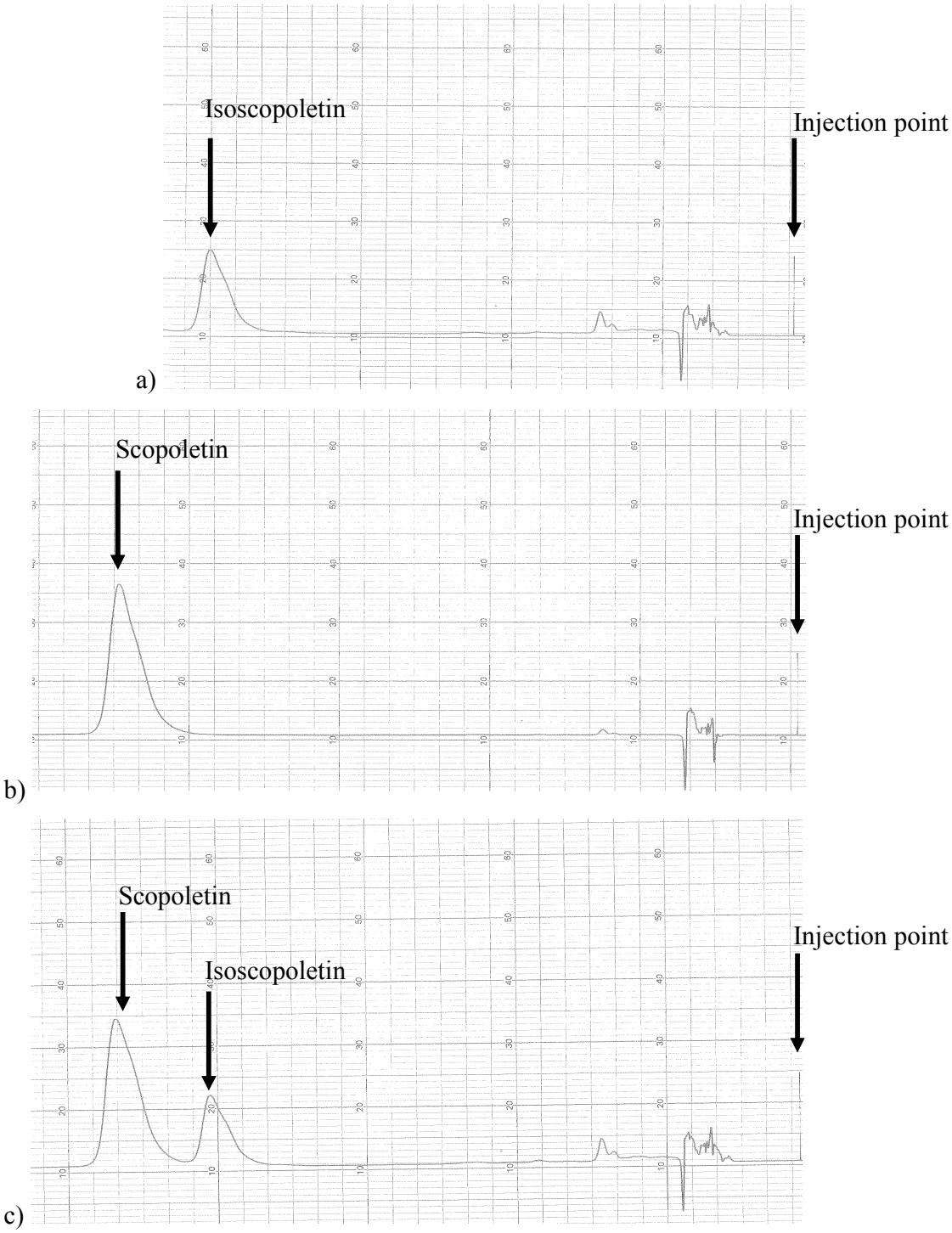


Fig. 2.8. HPLC traces of a) isoscopoletin, b) scopoletin, and c) their co-injection

Compound (2) (40 mg) was identified as linamarin. ^{1}H NMR showed the anomeric proton at δ 4.63 which is doublet, J = 7.8 Hz confirming the β-configuration of the sugar in addition to other protons confirming the structure. The ^{13}C NMR showed 10 signals corresponding to the 10 carbons: two methyl, one methylene, five methane and two quaternary carbons. Linamarin was originally identified in cassava and the sugar configuration was confirmed by synthesis of isolinamarin in which the sugar was in α-configuration and their properties and NMR data were compared (Clapp et al., 1966).

Compound (3) was identified as esculin by comparison of UV and HPLC retention time at 6.0 min with an authentic sample.

Compound (4) (7 mg) was identified as scopolin. In the ^{1}H NMR spectrum, the anomeric proton at δ5.09 J = 8 confirmed the β-configuration of glucose in addition to other sugar protons (Fig. 2.9, 2.10), ^{13}C NMR (Fig. 2.9, 2.11) and HR MS were as previously reported (Fliniaux et al., 1997).

Compound (5) was identified as esculetin by comparison of UV and HPLC retention time at 11.5 min with an authentic sample.

Compounds identified from deteriorated cassava root extract were scopoletin, linamarin, esculin, scopolin and esculetin are summarised in Fig 2.12.

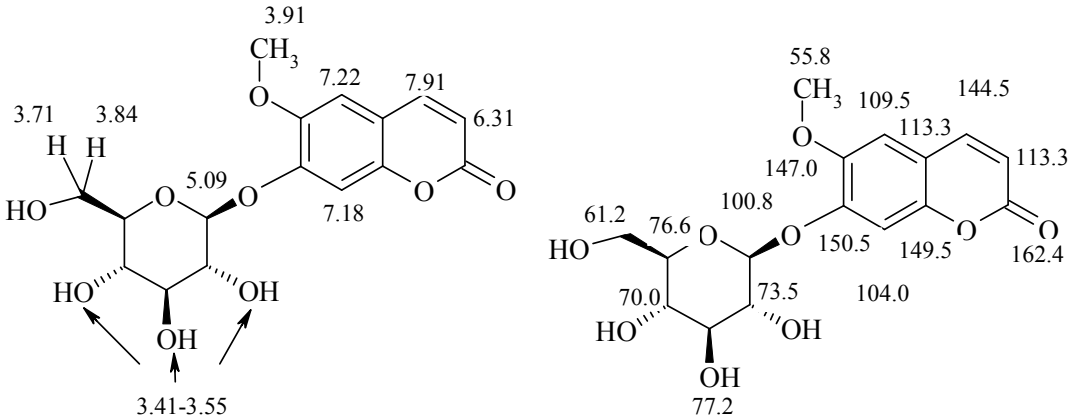


Fig. 2.9. scopolin (4) showing its ^{1}H NMR and ^{13}C NMR data.

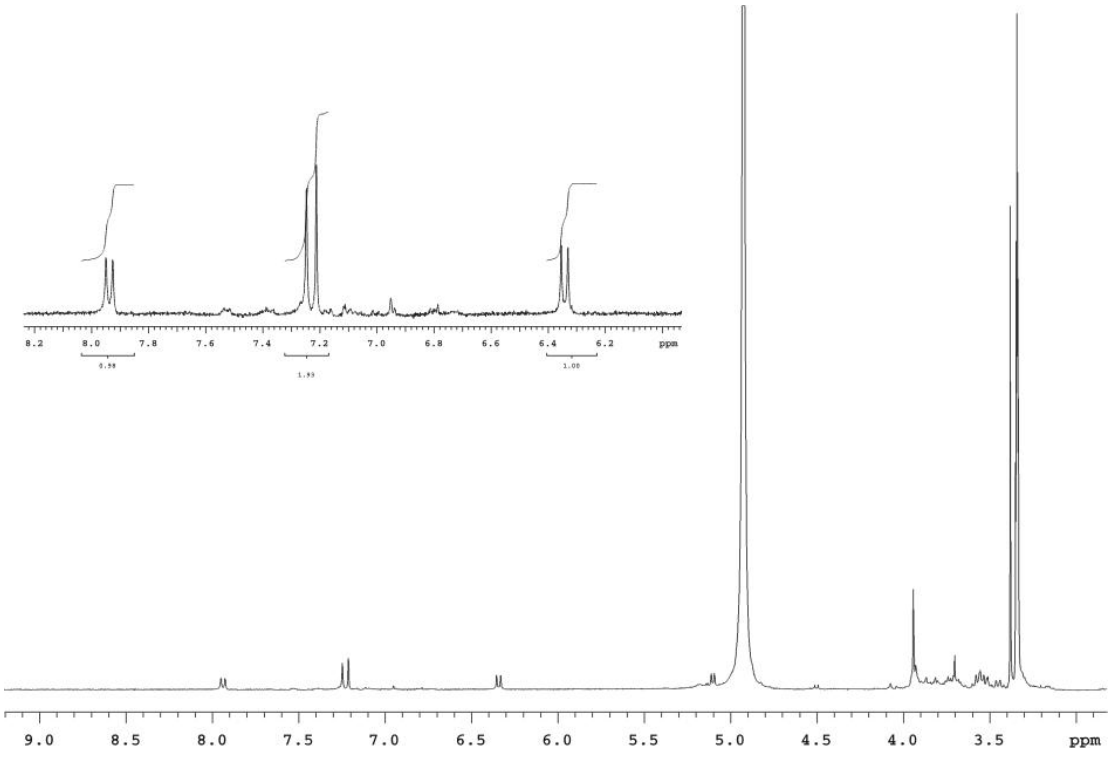


Fig. 2.10. ^{1}H NMR spectrum (CD_3OD, 400 MHz) of scopolin (4).

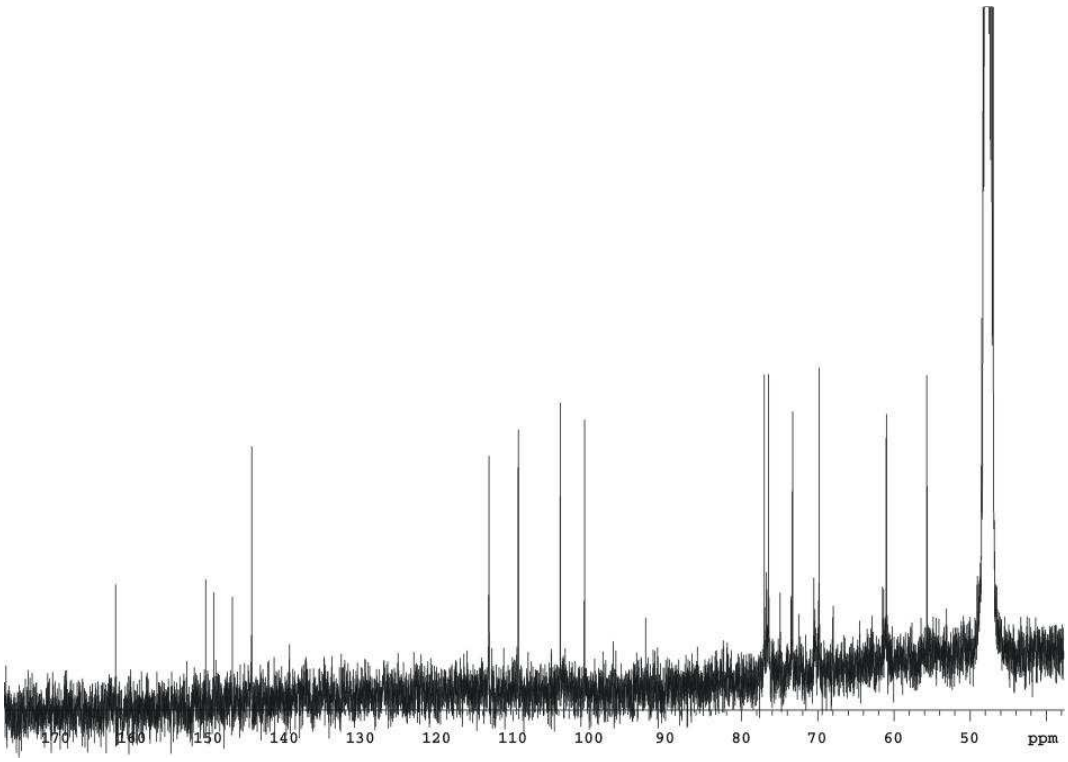
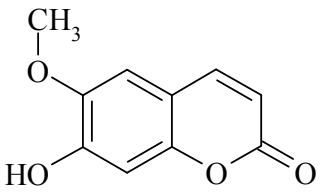
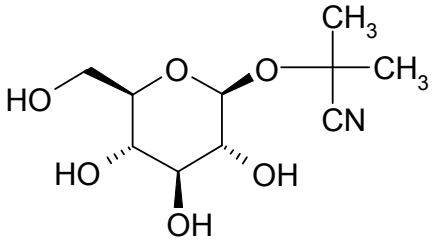


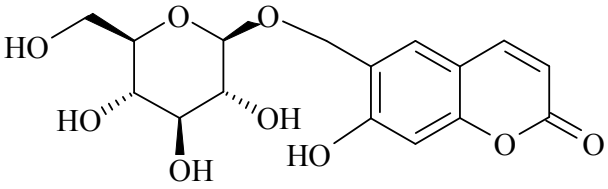
Fig. 2.11. ^{13}C NMR spectrum (CD_3OD, 400 MHz) of scopolin (4).



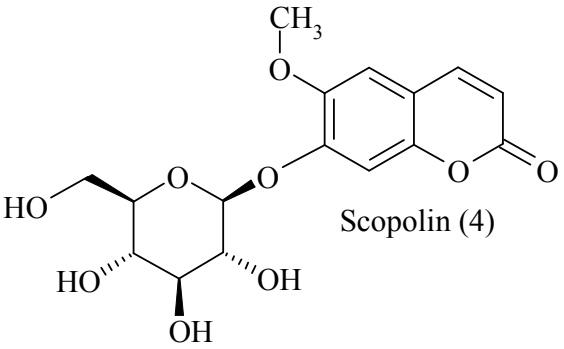
Scopoletin (1)



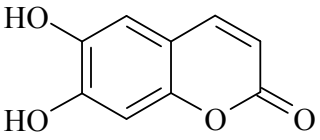
linamarin (2)



Esculin (3)



Scopolin (4)



Esculetin (5)

Fig. 2.12. Compounds identified from deteriorated cassava root extract.

Investigation of constituents in healthy cassava root

β-Carotene (6)

Compound (6) was identified by UV as β-carotene (Fig. 2.13) UV λ_{max} (EtOAc): 430, 452, and 475 nm. β-Carotene has previously been identified in six cassava cultivars. Leaves contained 13-78 mg/kg fresh weight while root contained 0.1-3 mg/kg fresh weight of β-carotene (Adewusi and Bradbury, 1993).

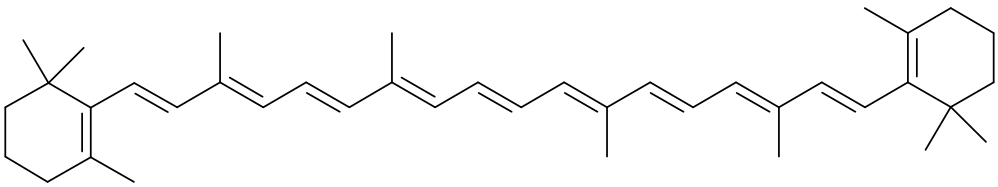


Fig. 2.13. β-Carotene (6).

Glycoside of diacylglyceride (7)

Compound (7) (10 mg per 5 kg fresh weight) was identified as a glycoside of diacylglyceride (Fig. 2.14) by NMR spectroscopy (Fig. 2.15), MS: ^{1}H NMR ($CDCl_3$-CD_3OD 9:1v/v) shows diagnostic signals of glycerol and galactose. One anomeric signal (H-1, doublet, 4.2 ppm) and its coupling constant (J = 6.9 Hz) indicated β-configuration of the galactosyl residue. In addition, ^{13}C NMR ($CDCl_3$-CD_3OD 9:1v/v) confirmed the structure of the sugar as galactose: δ 103.9 (C-1), 70.1 (C-2), 74.5 (C-3), 68.2 (C-4), 73.12 (C-5), 61.0 (C-6). The sample was not completely soluble in ($CDCl_3$-CD_3OD 9:1v/v) which caused broadness in the peaks, so ^{1}H NMR was repeated in pyridine-d_5 (400 MHz) and ^{13}C NMR (100 MHz) spectral data are summarised in Table 2.1. The anomeric protons and other signals were shifted when pyridine-d_5 was used in comparison with using $CDCl_3$-CD_3OD 9:1v/v as solvent. Spectral data are compatible with previous published data (Orgambide et al., 1992; Schröder et al., 2003; Janwitayanuchit et al., 2003).

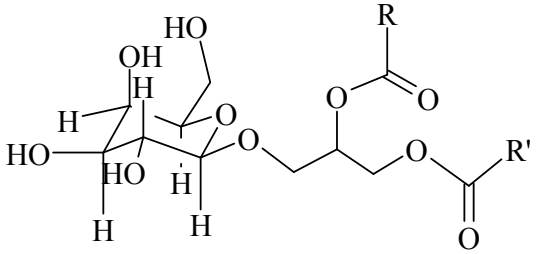


Fig. 2.14. Galactoside of diacyl diacylglyceride (7).

β-Configuration of the sugar residue was confirmed due to the large J constant of the anomeric proton at 4.81, d, J = 8 and the sugar was confirmed as galactose because H-3' at 4.14, dd, J = 3, 9 Hz show a small and a large coupling constant due to axial equatorial coupling (H-3', H4') and axial, axial coupling (H-3', H-2'). H-4'at 4.55, d, J = 3 Hz also showed this small coupling constant due to axial equatorial coupling (H-3', H-4').

The presence of six olefinic carbon signals in the ^{13}C NMR spectrum together with one signal for a methylene atom between two double bonds indicated that one of the fatty acid was mono-unsaturated and one was a methylene interrupted diene (CH=CH-**CH_2**-CH=CH).

Table 2.1

^{1}H NMR (400 MHz) and ^{13}C NMR (100 MHz) spectral data of compound (7) (pyridine-d_5)

Position	**δH multiplicity *J* (Hz)**	**δ C (ppm)**
Fatty acyl moiety		
Carbonyl	-	173.1, 173.3
α-Methylene	2.30 – 2.33, m	34.1, 34.4
β-Methylene	1.60, m	25.2*
Bulk methylene	1.23 – 1.33, m	22.8 – 32.1
Allylic methylene (**CH_2**-CH=CH)	2.05 -2.09, m	27.4, 27.5
Diallylic methylene (CH=CH-**CH_2**-CH=CH)	2.89, t, *J* = 6	26.0
Olefinic	5.43 – 5.52, m	128.3, 128.4, 130.1, 130.2, 130.3, 130.4
Methyl	0.82 – 0.86, m*	14.2, 14.3
Glycerol moiety		
H-1a	4.67, dd, *J* = 5, 12	63.2
H-1b	4.50, dd, *J* = 7, 12	
H-2	5.64 – 5.66, m	70.9
H-3a	4.05, m*	
H-3b	4.35, dd, *J* = 5, 11	68.0
Sugar moiety		
1'	4.81, d, *J* = 8	105.7
2'	4.43, d, *J* = 9	72.2
3'	4.14, dd, *J* = 3, 9	75.2
4'	4.55, d, *J* = 3	70.0
5'	4.05, t, *J* = 6*	77.1
6'	4.43, m*	62.2

* Overlapped

FAB-MS shows the presence of two molecular species differing by two mass units. The higher mass component agrees with a molecular formula of $C_{45}H_{82}O_{10}$, which would contain only two double bonds; either two mono-unsaturated acyl groups or one di-unsaturated acyl group and one saturated. The lower mass component agrees with a molecular formula of $C_{45}H_{80}O_{10}$ which would contain three double bonds.

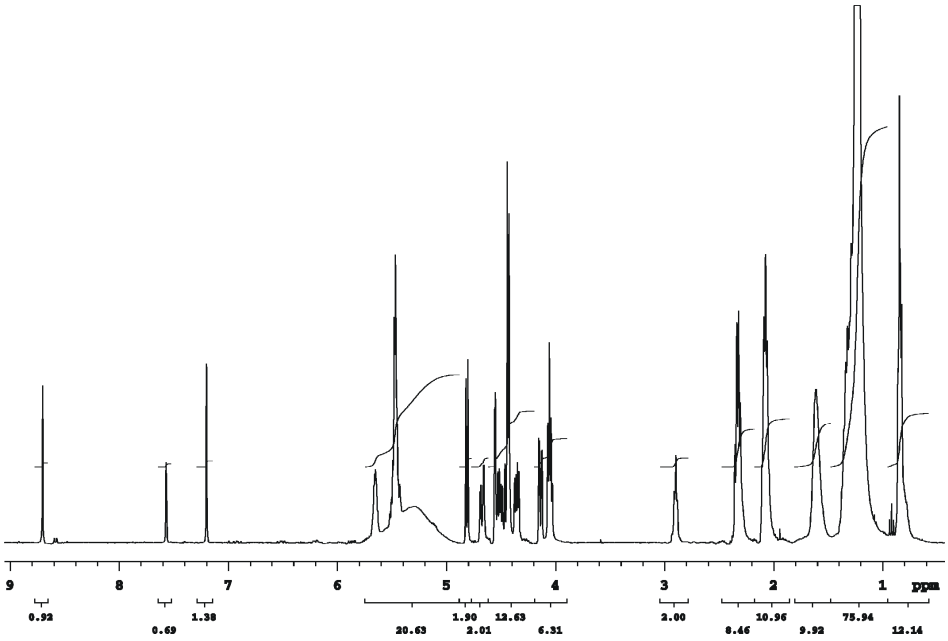


Fig. 2.15. ^{1}H NMR spectrum (pyridine-d_5, 400 MHz) of galactoside diacylglyceride (7).

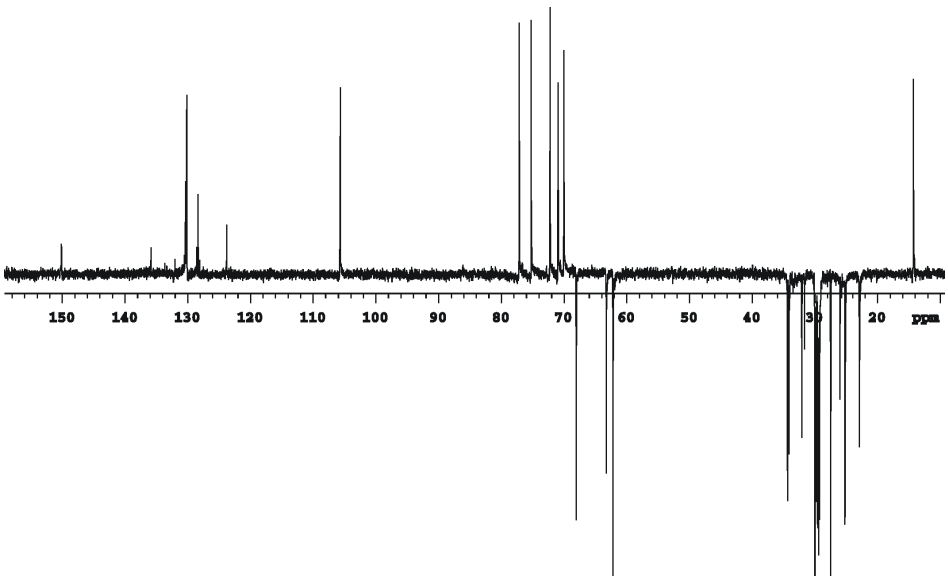


Fig. 2.16. DEPT spectrum (pyridine-d_5, 100 MHz) of galactoside diacylglyceride (7).

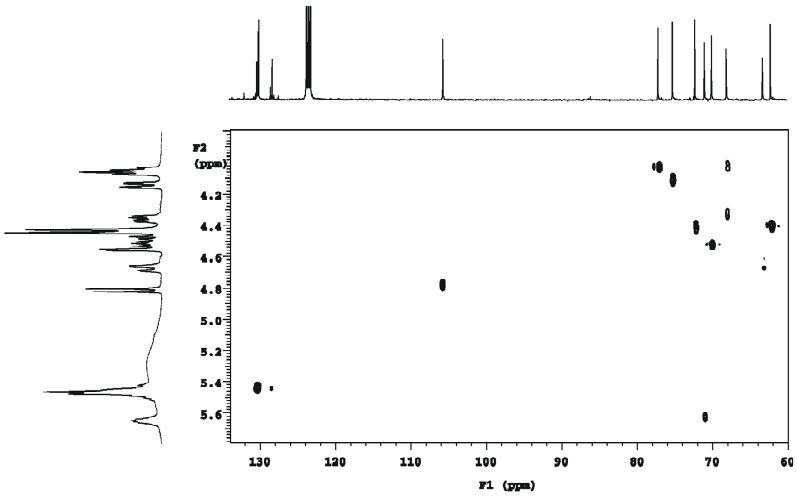


Fig. 2.17. Part of HMQC spectrum (pyridine-d_5, ^{1}H at 400, ^{13}C at 100 MHz) of galactoside diacylglyceride (7), showing correlations in the sugar and glycerol regions of the spectrum.

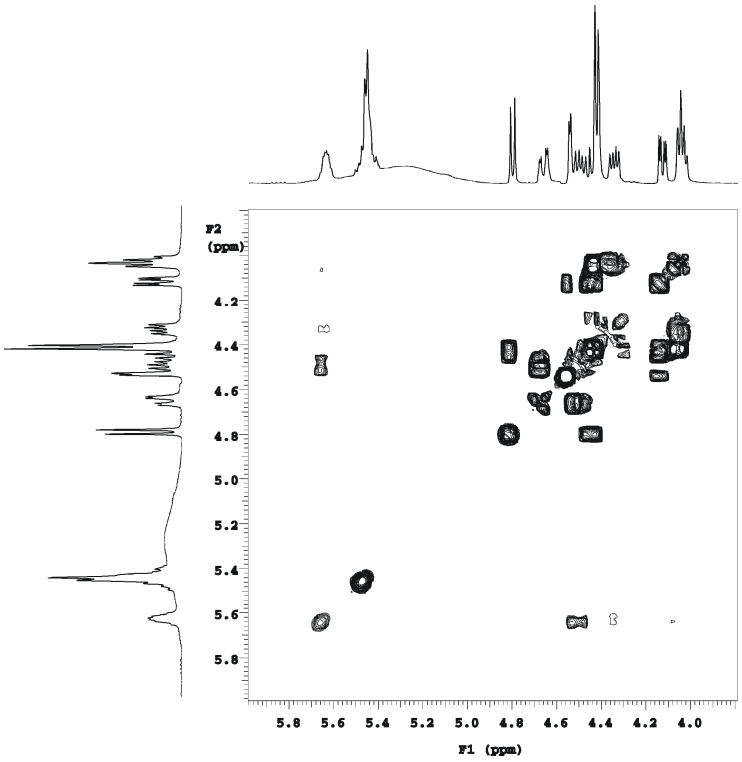


Fig. 2.18. Part of COSY spectrum (pyridine-d_5, 400 MHz) of galactoside diacylglyceride (7), showing correlations in the sugar and glycerol regions of the spectrum.

DEPT (Fig. 2.16), HMQC (Fig. 2.17), COSY (Fig. 2.18), and NOESY (not shown) spectral data confirmed these results. In HMQC carbons C-1 at δ 63.2 of the glycerol showed correlation with two protons H-1a at δ 4.67, dd, *J* = 5, 12 Hz and H-1b at δ 4.50, dd, *J* = 7, 12 Hz. The deshielded position of these protons relative to the other $-CH_2-$ protons of the glyceryl moiety suggests that an acyl group is attached this position. The carbon atom C-3 of the glycerol at δ 68.0 showed correlation with two protons H-3a at δ 4.05, m and H-3b at δ 4.35, dd, *J* = 5, 11 Hz suggesting that the ether linkage to galactose is at this position.

COSY spectral data (Fig. 2.18) confirmed the structure; showing correlation between allylic methylene (**CH_2**-CH=CH) at δ 2.05-2.09, m, diallylic methylene (CH=CH-**CH_2**-CH=CH) at δ 2.89, t, *J* = 6 Hz and olefinic protons at δ 5.43-5.52, m. There was correlation between α-methylene at δ 2.30 – 2.33, m and β-methylene at δ 1.60, m. The anomeric proton H-1' at 4.81, d, *J* = 8 Hz correlated to H-2' at 4.43, d, *J* = 9 Hz and also H-2' correlated to H-3' at 4.14, dd, *J* = 3, 9 Hz. H-3' correlated with H-2' and H-4' at 4.55, d, *J* = 3 Hz. H-6' at 4.43, m was overlapped with H-2' correlated with H-5' at 4.05, t, *J* = 6 Hz. NOESY spectral data confirmed the position of galactose on C-3 of the glycerol moiety by showing correlation between H-1' at 4.81and H-3a at 4.05.

Fig. 2.19. Galactoside of diacylglycerides (7a) (R = is mono-unsaturated acyl, R' is mono- or di-unsaturated acyl group in the two molecules) and cerebroside (7b)

There is some structural similarity (Fig. 2.19) between glycoside diacylglyceride (Schröder et al., 2003) and cerebroside; the NMR data of cerebroside have been published (Bruzik and Nyholm, 1997). Glycoside diacylglycerides are components of plastid membranes and many plastids are present in starch-containing organs (Kojima et al., 1991; Klaus et al., 2002). Biological activities previously reported for these compounds include antitumor-promoting, oxygen scavenging and anti-herpes simplex virus activity (Janwitayanuchit et al., 2003). This is the first time it has been described from cassava.

β-Sitosterol glucoside (8)

β-Sitosterol glucoside (8) (20 mg) was identified by ^{1}H NMR and ^{13}C NMR (pyridine-d_5) (Tables 2.2, 2.3) (Fig. 2.20, 2.21) and comparison with previous published data (Kojima et al., 1990).

Table 2.2

^{1}H NMR (400 MHz) and the correlated ^{13}C NMR (100 MHz) spectral data (pyridine-d_5)

position	**δH multiplicity *J* (Hz)**	**Correlated carbons (one bond) δ (ppm)**
3	3.91-3.99 m*	78.1
6	5.33 m	121.9
Me-18	0.63,s	11.95
Me-19	0.91,s	19.2
Me-21	0.96 d (7)	19.0
Me-26	0.83 - 0.90*	20.0
Me-27	0.83 - 0.90*	19.4
Me-29	0.83 - 0.90*	12.1
1'	5.04 d (8)	102.6
2'	4.05 t (8)	75.3
3'	4.28 m*	78.6
4'	4.28 m*	71.7
5'	3.91-3.99 m*	78.5
6'a	4.55 dd (12, 2)	62.8
6'b	4.39 dd (12, 5)	62.8

* Overlapped signals

Table 2.3

^{13}C NMR (100 MHz) spectral data

position	**δ (ppm)**	**position**	**δ (ppm)**
1	37.5	19	19.2
2	30.2	20	36.4
3	78.1	21	19.0
4	39.3	22	34.2
5	140.9	23	26.4
6	121.9	24	46.0
7	32.2	25	29.4
8	32.0	26	20.0
9	50.3	27	19.4
10	36.9	28	23.4
11	21.3	29	12.1
12	39.9	1'	102.6
13	42.5	2'	75.3
14	56.8	3'	78.6
15	24.5	4'	71.7
16	28.5	5'	78.5
17	56.2	6'	62.8
18	12.0		

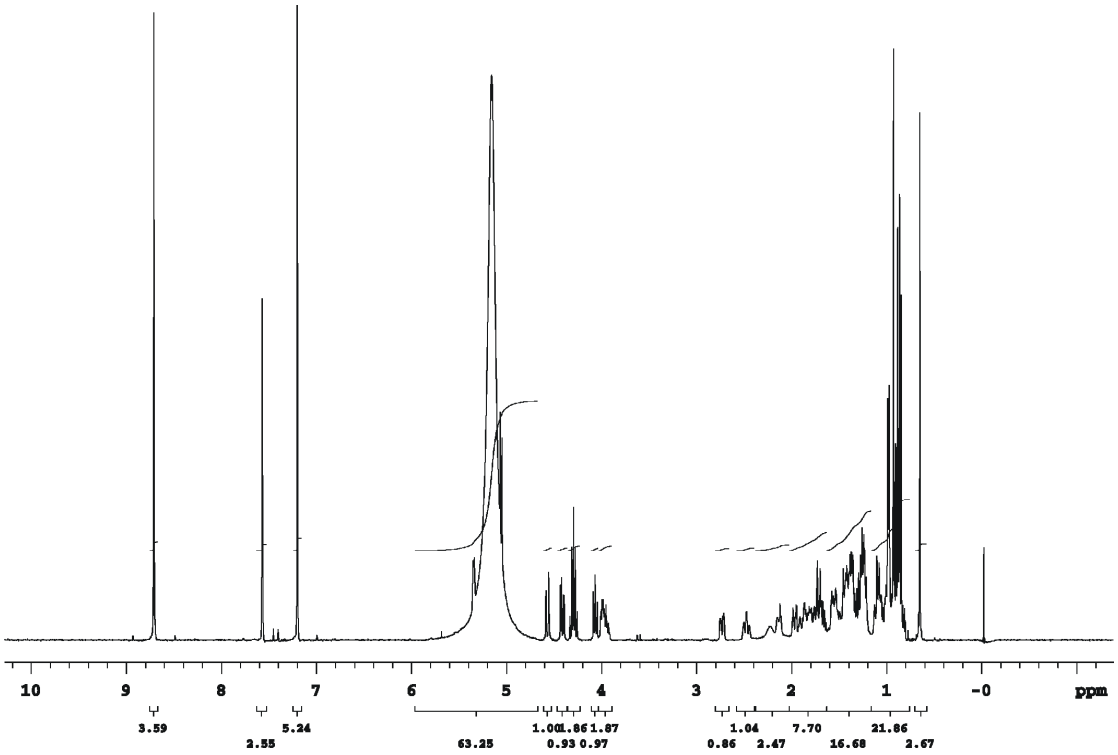


Fig. 2.20. ^{1}H NMR spectrum (pyridine-d_5, 400 MHz) of β-sitosterol glucoside.

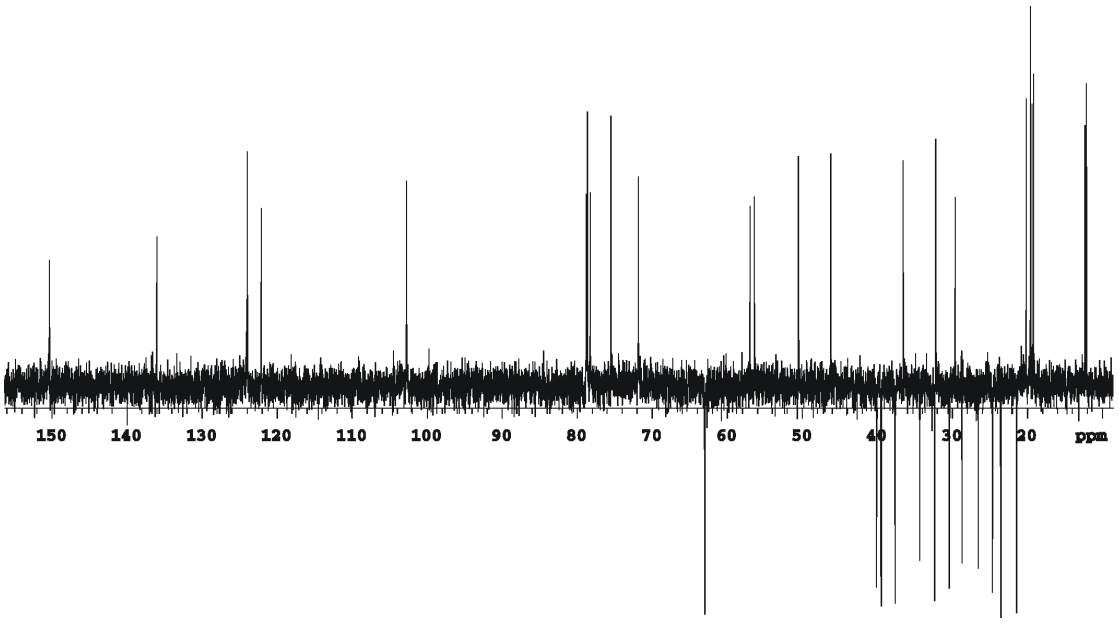


Fig. 2.21. DEPT spectrum (pyridine-d_5, 100 MHz) of β-sitosterol glucoside.

The HMQC spectrum showed correlation between the anomeric proton at δ 5.04 and its carbon at δ 102.6, olefinic proton δ 5.33and its carbon at δ 121.9, the two overlapped protons at 4.28 related to C-3' and 4' at 78.6 and 71.7, while the two overlapped protons at 3.91-3.99 related to C-5' and 3 at 78.5 and 78.1, H-2' at 4.05 correlated to its C-2' at 75.3, H-6'a at 4.55 and H-6'b at 4.39 related to C-6' at 62.8, the overlapped methyl groups at 0.83 - 0.90 are correlated to C-26, 27 and 29 at 20.0, 19.4 and 12.1; methyl at 0.63 correlated to C-18 at 11.95 and methyl at 0.91 correlated to C-19 at 19.2. HR ESI MS confirmed our results: *m/z* of $[M+NH_4]^+$ $C_{35}H_{64}O_6N_1$ requires 594.4728, found 594.4733.

Compound (8) was identified as β-sitosterol glucoside (Fig. 2.22). It is one of the phytosterol glucosides that are present in many plants (Rizk and Rimpler, 1972; Kojima et al., 1990; Khan et al., 1991). It participates in cellulose synthesis (Peng et al., 2002); this is the first time that it has been identified from cassava.

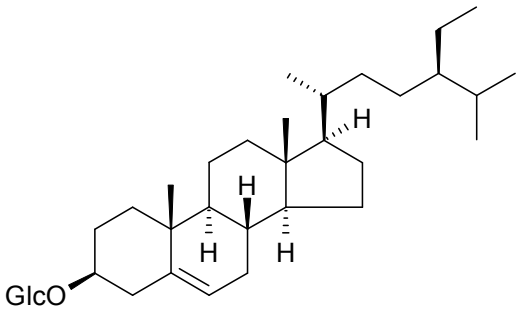
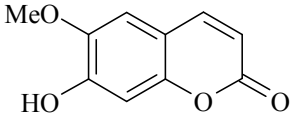
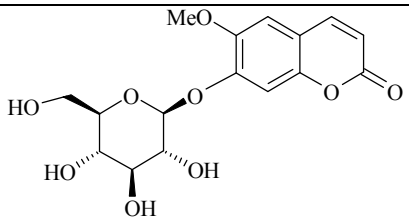
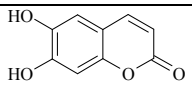
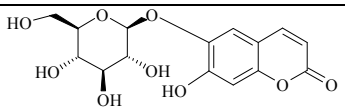
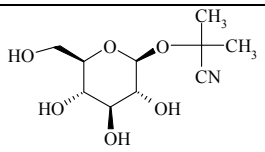
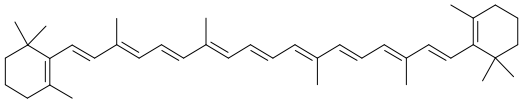
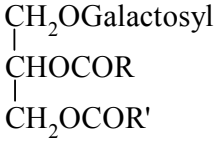
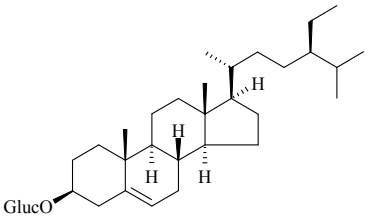


Fig. 2.22. β-Sitosterol glucoside (8).

The dramatic accumulation of pharmacologically active hydroxycoumarins (mainly scopoletin and scopolin) in cassava roots during PPD may play roles as antioxidants and in plant defence. In this report the rapid increase in scopoletin and scopolin accumulation in cassava roots during post-harvest deterioration provided an ideal system with which to investigate alternative pathways for the biosynthesis of these hydroxycoumarins using different feeding and competition experiments. Despite their importance, key aspects of the biosynthesis of these secondary metabolites remain unresolved in plants including cassava.

Table 2.4

Compounds that were isolated and identified in this study from (healthy and deteriorated) cassava roots

Name	**Structure**	**Weight/fresh weight**
Scopoletin		3 mg / 1 kg
Scopolin		5 mg / 1 kg
Esculetin		less than 1 mg / 1 kg
Esculin		less than 1 mg / 1 kg
Linamarin		29 mg / 1 kg
β-Carotene		less than 1 mg / 1 kg
Galactoside of diacylglyceride		2 mg / 1 kg
β-Sitosterol glucoside		4 mg / 1 kg

Chapter 3

Biosynthesis of hydroxycoumarins in cassava roots during PPD: the *E*-*Z*-isomerisation stage

3.1. Introduction

Cassava (*Manihot esculenta* Crantz Family Euphorbiaceae) is an important economical and nutritional starch-rich crop, relatively easy to grow even in poor soils and under drought conditions. The occurrence after only one to three days of blue to black vascular streaking, post-harvest physiological deterioration (PPD), causes significant wastage and economic losses. PPD has been explained as a physiological process not due to microorganisms (Averre, 1967; Noon and Booth, 1977) and on a molecular basis as an oxidative burst which initiates within 15 min of the root being injured, (Reilly et al., 2003; Reilly et al., 2004) followed by altered gene regulation (Reilly et al., 2007). The accumulation of secondary metabolites is important in PPD (Tanaka et al., 1983; Buschmann et al., 2000b). Amongst these secondary metabolites are hydroxycoumarins (e.g. scopoletin and its glucoside scopolin) which show antioxidant properties and which may by oxidation and polymerisation give rise to the blue/black discolouration. These hydroxycoumarins may be important in plant defence as phytoalexins due to the induction of their biosynthesis following various stress events (wounding, bacterial and fungal infections) (Giesemann et al., 1986; Gutierrez et al., 1995). However, their biosynthesis in cassava is not known and neither is it clearly understood in other plants (Petersen et al., 1999).

As part of on-going studies into these hydroxycoumarins (Buschmann et al., 2000b; Reilly et al., 2003; Reilly et al., 2004), the incorporation of cinnamic acid-d_7 into cassava roots under PPD was investigated, in order to confirm that it was a potential precursor of scopoletin and then of scopolin on the phenylpropanoid pathway (see Chapter 1), from phenylalanine following the action of phenylalanine ammonia lyase (PAL). The unexpected experimental result that scopoletin-d_3 was produced, and not scopoletin-d_4 (Fig. 3.1), has led us to study in detail the *E*-*Z*-isomerisation step in the biosynthesis of scopoletin. This is the most likely biosynthetic step at which loss of an additional deuterium atom might occur. Furthermore, this isomerisation is not resolved in plants, and has not been previously reported in cassava under PPD. Here we exploit the increase in hydroxycoumarin accumulation in cassava roots post-harvest to investigate this isomerisation step in the biosynthesis of scopoletin.

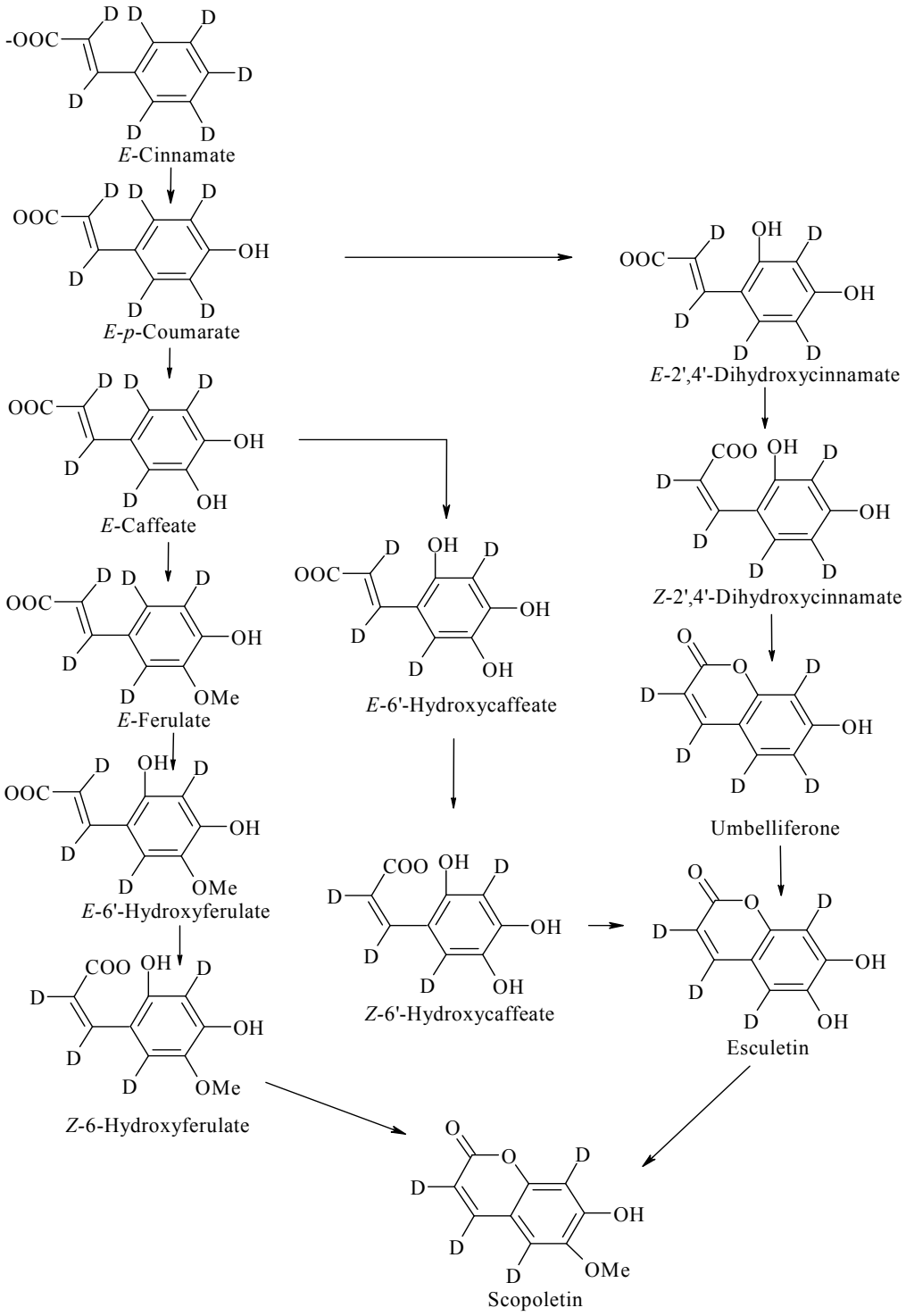


Fig. 3.1. Proposed metabolic pathways of the biosynthesis of scopoletin after feeding cassava roots with *E*-cinnamic-d_7 acid showing the expected scopoletin-d_4. An *E*-*Z* isomerisation of one of the cinnamic acid derivatives is required to allow the formation of the lactone ring.

3.2. Experimental

Plant material

Root tubers of different cultivars (MCOL 22, MNGA 19, MNGA 2) were harvested from cassava plants growing in the tropical glass house at the University of Bath as given above in Chapter 2.

General methods

Chemicals were obtained routinely from Sigma-Aldrich Chemical Co. Ltd, UK. *E*-Cinnamic-2',3',4',5',6'-d_5 acid was obtained from CDN Isotopes, Canada. All solvents used were supplied by Fisher, all in GPR grade, except where stated. The HPLC instrument consisted of a solvent delivery system equipped with Jasco PU-980 pump and monitored at 360 nm with a Jasco UV-975 detector, using 16% acetonitrile in water containing 0.1% formic acid, flow rate 4 ml/min at 20 °C. The chromatograms were recorded on a Goerz Metrawatt Servogor 120 recorder. HPLC columns were purchased from Phenomenex Inc.: Phenomenex Gemini 10 μ C18 110A 250 x 10 mm with guard column Phenomenex Gemini 5 μ C18 10×10 mm**.** Samples were injected using a 100 μl loop. HR ESI MS was carried out on a Bruker micrOTOF mass spectrometer in the department of Pharmacy and Pharmacology, University of Bath or on a Micromass Quattro II in EPSRC National Mass Spectrometry Service Centre, University of Wales, Swansea. NMR spectra were obtained on a Varian Mercury Spectrometer at 400 MHz (^{1}H) and 100.8 MHz (^{13}C) in CD_3OD, all chemical shifts are reported in parts per million (ppm) relative to internal tetramethylsilane, and coupling constants (*J* values, line separations) are absolute values in Hz.

Extraction, purification and characterisation of scopolin and scopoletin

Cassava roots (1.3 kg, various cultivars) were peeled (1.1 kg) and then cut into approximately 1 cm^3 cubes, allowed to deteriorate for up to 6 d under controlled conditions (20° C, 80-90% R.H.) until crushing and extraction. Daily from day 3, approx. 270 g of the sliced roots were crushed and macerated (EtOH 200 ml). The EtOH extracts of these 4 samples were analysed by TLC ($CHCl_3$: EtOAc: MeOH 2: 2:

1,visualised by UV at 365 nm) and there was no significant difference between the first two samples, days 3 and 4, where it is known that the highest amount of scopoletin and scopolin accumulation occurs (Buschmann et al., 2000b). The combined ethanolic extracts were filtered and evaporated under reduced pressure at 35-40 °C. The residue was fractionated between water (see below) and $CHCl_3$. The combined $CHCl_3$ layers were concentrated to yield a pale yellow viscous oil (240 mg) which was then further purified by solid-phase extraction to afford scopoletin as given above in Chapter 2, identical by TLC to a commercial sample. The residue was dissolved in MeOH (0.2 g/ml) and purified by HPLC. A 24.7 min peak was detected, collected, concentrated and dissolved in MeOH and identified by high field NMR spectroscopy and HR ESI MS (Hirata et al., 2000) as the coumarin, scopoletin (4 mg). ^{1}H NMR (CD_3OD) is given above in Chapter 2.

The concentrated aq. fraction (7.6 g) was separated using the above HPLC conditions. A 7.1 (and a trace of a 27.4) min peak was detected, collected, concentrated and dissolved in MeOH and identified by high field NMR spectroscopy and HR MS (Fliniaux et al., 1997) as the coumarin, scopolin (7 mg). ^{1}H NMR (CD_3OD) is given above in Chapter 2.

General feeding procedure

Cassava roots (typically 1 kg) were peeled, then cut into approximately 1 cm^3 cubes and divided into groups (typically 100 g). One group was immediately crushed and extracted with EtOH (fresh cassava extract) and another (control group) was stored under controlled conditions (20° C, 80-90% R.H.) until crushing and extraction. Where potential biosynthetic precursors were introduced, this was by spraying a group of cubes with the precursor in solution with a simple hand-pumped aerosoliser. In this series of experiments, labelled cinnamic acid (typically 30 mg) was dissolved in aq. 4% Na_2CO_3 (3 ml) then adjusted to pH 7.5 with 1 M HCl. As the highest amount of scopoletin and scopolin accumulation occurs between days 3 and 4 (this is cultivar dependent (Buschmann et al., 2000b)), after 3 d, half the group was crushed and macerated (EtOH 200 ml, 3 d), and the other half was crushed after 4 d and then macerated (EtOH 200 ml, 2 d). The combined ethanolic extracts were filtered. A second round of maceration (EtOH, 400 ml, 2 d) was

performed. The combined EtOH extracts (after 8 d), were filtered and evaporated under reduced pressure at 35-40 °C. The residue was dissolved in MeOH (0.2 g/ml) and purified by HPLC. From repeat injections (n = 12, each of 100 µl) two peaks of retention time 7.1 min and 24.7 min (scopolin and scopoletin respectively, the latter by comparison with a commercial authentic sample) were detected, collected, concentrated and dissolved in MeOH and identified by High Resolution Electrospray Ionisation Mass Spectrometry (HR ESI MS) in order to separate isotopes and to identify the percentage of labelled scopoletin and scopolin in relation to the unlabelled coumarins.

Feeding with E-cinnamic-d_7 acid

Using the General Feeding procedure, cassava roots (1.4 kg, cv MNGA 2) were peeled (1.2 kg) and fed with *E*-cinnamic-2,3,2',3',4',5',6'-d_7 acid (150 mg) dissolved in aq. 4% Na_2CO_3 (10 ml). A representative sample of the combined EtOH extract (32.9 g) was then purified by HPLC. The HPLC peak at 7.1 min, was consistent with scopolin, naturally occurring scopolin $C_{16}H_{19}O_9$ requires 355.1024, HR MS found *m/z* 355.1041 $[M + H]^+$, $C_{16}H_{18}O_9Na$ requires 377.0843, found 377.0847 $[M + Na]^+$, also $C_{16}H_{16}D_3O_9$ requires 358.1212, found 358.1229 $[M + H]^+$, and $C_{16}H_{15}D_3O_9Na$ requires 380.1031, found 380.1017 $[M + Na]^+$.

The HPLC peak at 24.7 min, was consisted with scopoletin; naturally occurring scopoletin $C_{10}H_9O_4$ requires 193.0495, HR MS found *m/z* 193.0497 $[M + H]^+$, $C_{10}H_8O_4Na$ requires 215.0315, found 215.0309 $[M + Na]^+$, also $C_{10}H_6D_3O_4$ requires 196.0684, found 196.0687 $[M + H]^+$, $C_{10}H_5D_3O_4Na$ requires 218.0503, found 218.0506 $[M + Na]^+$, $C_{20}H_{16}O_8Na$ requires 407.0737, found 407.0729 $[2M + Na]^+$, and $C_{20}H_{13}D_3O_8Na$ requires 410.0921, found 410.0926 $[2M + Na]^+$.

Feeding with E-cinnamic-d_6 acid

Using the General Feeding procedure, cassava roots (1.23 kg, cv MCOL 22) were peeled (1 kg) and a group (195 g) was fed with *E*-cinnamic-3,2',3',4',5',6'-d_6 acid (30 mg) dissolved in aq. 4% Na_2CO_3 (3 ml). A representative sample of the combined EtOH extract (~4 g) was then purified by HPLC. The HPLC peak at 7.1 min, scopolin; $C_{16}H_{19}O_9$ requires 355.1024, HR MS *m/z* found 355.1037 $[M + H]^+$, $C_{16}H_{18}O_9Na$

requires 377.0843, found 377.0852 $[M + Na]^+$, also $C_{16}H_{16}D_3O_9$ requires 358.1212, found 358.1215 $[M + H]^+$ and $C_{16}H_{15}D_3O_9Na$ requires 380.1031, found 380.1029 $[M + Na]^+$.

The HPLC peak at 24.7 min, scopoletin; naturally occurring scopoletin $C_{10}H_9O_4$ requires 193.0495, HR MS *m/z* found 193.0488 $[M + H]^+$, $C_{10}H_8O_4Na$ requires 215.0315, found 215.0314 $[M + Na]^+$, also $C_{10}H_6D_3O_4$ requires 196.0684, found 196.0689 $[M + H]^+$.

Synthesis of E-cinnamic-2-d_1 acid

Benzaldehyde (127 mg, 1.2 mmol) and malonic-d_4 acid (99 atom% d) (276 mg, 2.63 mmol) were dissolved in pyridine (660 µl). Piperidine (13 µl, 131 nmol, 0.00011 eq.) was added and the reaction was heated to 68 ºC (oil bath) for 24 h. Water (7.5 ml) and then conc. HCl (0.4 ml) were added dropwise until a precipitate appeared which was collected and recrystallized (water) to afford a product, homogenous by TLC, $R_f = 0.46$, n-hexane-ethyl acetate-acetic acid (1:1:0.01, v/v/v), as white crystals 127-130 ºC (Lit. 133 ºC for unlabelled *E*-cinnamic acid, Merck Index), yield 59% relative to benzaldehyde used. $C_9H_6D_1O_2$ requires 148.0514, HR MS gave *m/z* 148.0516 $[M - H]^-$. ^{1}H NMR: δ 7.39 (3H, m, H-3',4',5'), 7.58 (2H, m, H-2',6'), 7.66 (0.88H, br s, H-3), as well as 6.48 (0.12H, d, $J = 16$) from the residual isotopomer at H-2 and 7.67 (0.12H, d, $J = 16$, H-3). ^{13}C NMR: δ 119.1 (C-2-D, 1:1:1 t, $^1J_{CD} = 25$), 129.2 (C-3',5'), 130.0 (C-2',6'), 131.4 (C-4'), 135.8 (C-1'), 146.2 (C-3), 170.3 (C-1), as well as a small singlet at 119.3 from the residual isotopomer at C-2. ^{2}H NMR (61.41 MHz): δ 6.51 (2-D, d, $J = 2$).

Feeding with E-cinnamic-2-d_1 acid, E-cinnamic-2',3',4',5',6'-d_5 acid, E-cinnamic-3,2',3',4',5',6'-d_6 acid and E-cinnamic-2,3,2',3',4',5',6'-d_7 acid

Using the General Feeding procedure, cassava roots (800 g, cv MCOL 22) were peeled (640 g) and divided into four groups (85 g) which were fed with *E*-cinnamic-2-d_1 acid, *E*-cinnamic-2',3',4',5',6'-d_5 acid, *E*-cinnamic-3,2',3',4',5',6'-d_6 acid and *E*-cinnamic-2,3,2',3',4',5',6'-d_7 acid (20 mg of each acid) dissolved in aq. 4% Na_2CO_3 (2.0 ml). A representative sample of the combined EtOH extract (~2 g) from each group was then purified by HPLC. The HPLC peak at 7.1 min scopolin, $C_{16}H_{19}O_9$ requires 355.1024,

HR MS *m/z* found 355.1033 $[M + H]^+$, and the corresponding HR MS data are listed in Table 3.1. The HPLC peak at 24.7 min, scopoletin; naturally occurring scopoletin $C_{10}H_9O_4$ requires 193.0495, HR MS *m/z* found 193.0489 $[M + H]^+$, and the corresponding HR MS data are listed in Table 3.2.

Feeding with E-cinnamic-2,3,2',3',4',5',6'-d_7 acid in light and in dark

Using the General Feeding procedure, cassava roots (690 g, cv MNGA 19) were peeled (490 g) and divided into two groups (85 g) which were fed with *E*-cinnamic-2,3,2',3',4',5',6'-d_7 acid (20 mg) dissolved in aq. 4% Na_2CO_3 (2.3 ml). One group was treated as above and then stored in the dark (in a box double wrapped with aluminium foil) while the other was treated as usual. Representative samples of the EtOH extract (~2 g) were then purified by HPLC. In the light conditions, the HPLC peak at 7.1 min, consistent with scopolin; naturally occurring scopolin $C_{16}H_{19}O_9$ requires 355.1024, HR MS found *m/z* 355.1068 $[M + H]^+$, $C_{16}H_{18}O_9Na$ requires 377.0843, found 377.0812 $[M + Na]^+$, also $C_{16}H_{16}D_3O_9$ requires 358.1212, found 358.1243 $[M + H]^+$, and $C_{16}H_{15}D_3O_9Na$ requires 380.1031, found 380.1021 $[M + Na]^+$. The percentage of scopolin-d_3 was 18.7%. In the dark conditions, naturally occurring scopolin $C_{16}H_{19}O_9$ requires 355.1024, HR MS found *m/z* 355.1049 $[M + H]^+$, $C_{16}H_{18}O_9Na$ requires 377.0843, found 377.0850 $[M + Na]^+$, also $C_{16}H_{16}D_3O_9$ requires 358.1212, found 358.1197 $[M + H]^+$, and $C_{16}H_{15}D_3O_9Na$ requires 380.1031, found 380.1069 $[M + Na]^+$. The percentage of scopolin-d_3 was 15.7%.

In the light conditions, the HPLC peak 24.7 min, scopoletin; naturally occurring scopoletin $C_{10}H_9O_4$ requires 193.0495, HR MS *m/z* found 193.0498 $[M + H]^+$, $C_{10}H_8O_4Na$ requires 215.0315, found 215.0324 $[M + Na]^+$, also $C_{10}H_6D_3O_4$ requires 196.0684, found 196.0703 $[M + H]^+$ and $C_{10}H_5D_3O_4Na$ requires 218.0503, found 218.0512 $[M + Na]^+$. The percentage of scopoletin-d_3 was 9.8%. In the dark conditions, the HPLC peak 24.7 min, scopoletin; $C_{10}H_9O_4$ requires 193.0495, HR MS *m/z* found 193.0494 $[M + H]^+$, $C_{10}H_8O_4Na$ requires 215.0315, found 215.0324 $[M + Na]^+$, also $C_{10}H_6D_3O_4$ requires 196.0684, found 196.0697 $[M + H]^+$ and $C_{10}H_5D_3O_4Na$ requires 218.0503, found 218.0527 $[M + Na]^+$. The percentage of scopoletin-d_3 was 9.1%.

3.2. Results and discussion

Hydroxycinnamates are utilized in various pathways in the formation of different phenylpropanoids including hydroxycoumarins (Fig. 3.1). *E*-Cinnamic acid is one of the early precursors in the biosynthesis of scopoletin under PPD (Fig. 3.1). An experiment was carried out feeding cubed cassava roots with deuterium labelled *E*-cinnamic-d_7 acid, in order to determine the level of incorporation of label observed under these simple conditions, and to investigate intermediates along the hydroxycoumarin biosynthetic pathway. However, the intriguing result that a scopoletin molecule labelled with only 3 deuterium atoms, scopoletin-d_3 and its *O*-glycoside scopolin-d_3 was obtained and not the (expected) corresponding d_4 analogues (Fig. 3.2) caused us to investigate the isomerisation stage in detail.

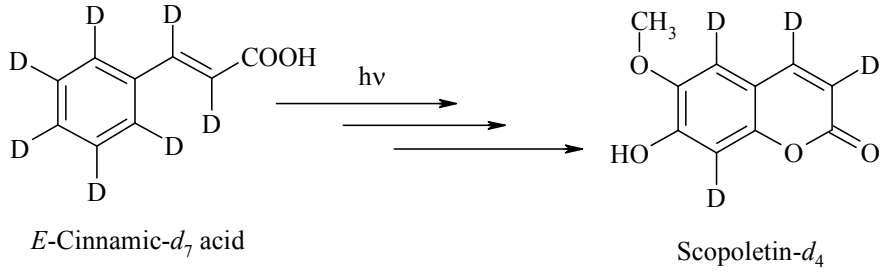


Fig. 3.2. The expected result of feeding *E*-cinnamic-d_7 acid is scopoletin-d_4.

Feeding experiments with deuteriated E-cinnamic acids

In order to investigate the biosynthesis of scopoletin and scopolin during PPD, a feeding experiment was developed. Freshly harvested cassava roots were fed with different deuteriated *E*-cinnamic acid intermediates, at pH 7.5 adjusted to be compatible with the physiological state of the roots. A commercial authentic sample of scopoletin was used to identify its HPLC peak in the root extracts, however, scopolin is not commercially available. Scopolin and scopoletin were collected and unambiguously identified by NMR and HR MS as described in chapter 2, and therefore our isolated scopolin was used as a reference sample.

E-cinnamic-2-d_1 acid is one required precursor for these isotopically labelled feeding experiments. Therefore *E*-cinnamic-2-d_1 acid was synthesised by a Knoevenagel reaction (Robbins and Schmidt, 2004; Ji et al., 2005), the nucleophilic addition of malonic acid-d_4 to the carbonyl group of benzaldehyde (Fig. 3.3), as in an aldol condensation, followed by decarboxylation step. The catalyst was a basic mixture of piperidine (0.00011 eq.) pK_a 11 and pyridine (as solvent) pK_a 5.5 (Clayden et al., 2001) at 68 ^oC for 24 h, and it afforded a white crystalline compound (59%, mp 127-130 ºC). *E*-cinnamic-2-d_1 acid has been synthesised previously from the corresponding α-trimethylated amino acid via Hofmann elimination in deuteriated water (Manitto et al., 1973), but no characterisation details were reported in the communication. HR MS gave *m/z* 148.0516 $[M - H]^-$, and comparison of the peak heights at *m/z* 148 (*E*-cinnamic-2-d_1 acid) and *m/z* 147 (unlabelled *E*-cinnamic acid) showed a deuterium label incorporation of 88%. Correspondingly, the ^{13}C NMR spectrum (Fig. 3.5) showed two key signals corresponding to C-2 for labelled and unlabelled cinnamic acid respectively: δ 119.1 (t, 1:1:1, $^1J_{CD}$ = 25 Hz), shifted upfield by 0.2 ppm by α-substitution with deuterium, typically 0.25 ppm per α-deuterium atom (Tulloch and Mazurek, 1973; Wehrli and Wirthlin, 1983; Hardick et al., 1996), and a small singlet at 119.3 for C-2 attached to proton, the residual isotopomer. In the ^1H NMR spectrum (Fig. 3.4), H-3 signals (integrating for 1) appear at 7.66 ppm as a slightly broadened singlet (incorporating a small $^3J_{HD}$ 2 Hz, integral 0.88) (Williams and Fleming, 2008b), superimposed on 7.67 (d, J = 16 Hz, integral 0.12) due to ^3J H-2–H-3 *E*-coupling. The small signal at δ 6.48 corresponding to H-2 (d, J = 16 Hz) due to ^3J H-2-H-3 *E*-coupling likewise integrated for 0.12, confirming the HR MS labelling result of 88%. The presence of the 12% of unlabelled cinnamic acid could be due to exchange of malonic acid-d_4 deuterium at position 2 with protons of piperidine. Finally, ^2H NMR observation (Fig. 3.6) confirmed these results by the signal at δ 6.51 (d, J = 2 Hz) which fits within the published range of 1-3 Hz for D-2-H-3 coupling as $J_{HH} = 6.5\ J_{HD}$ (Williams and Fleming, 2008). Also there is an upfield shift (0.03 ppm) which is less important than the ^{13}C chemical upfield shift (0.2 ppm) of C-2 which attached to D-2.

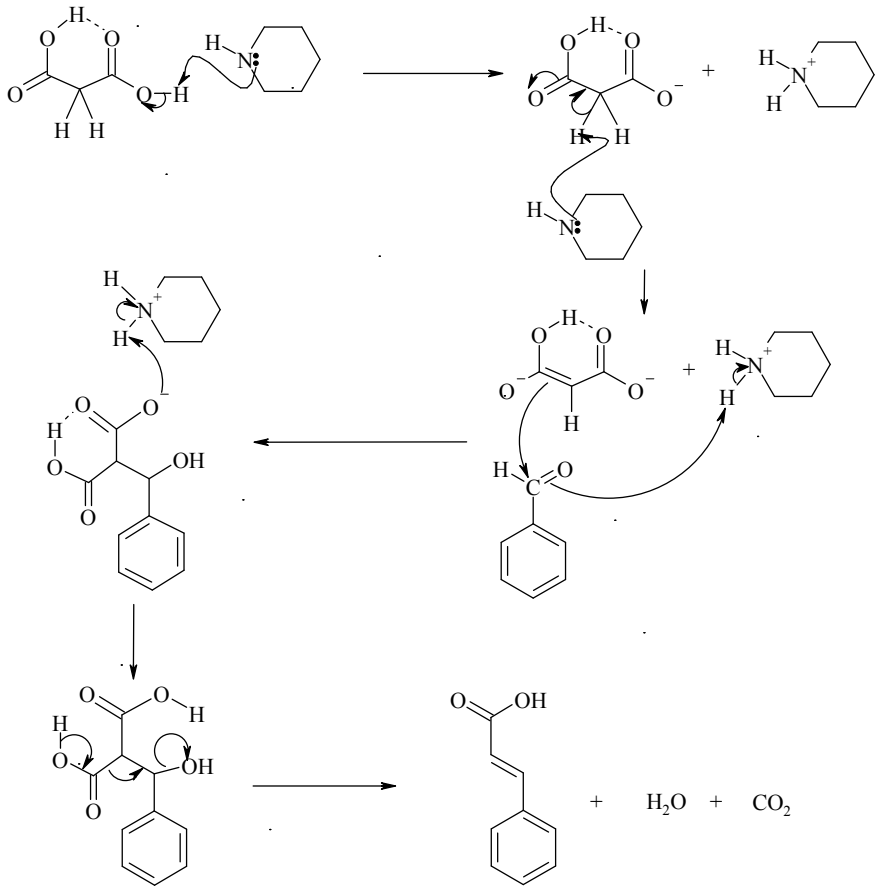


Fig. 3.3. Knoevenagel reaction for the synthesis of cinnamic acid derivatives.

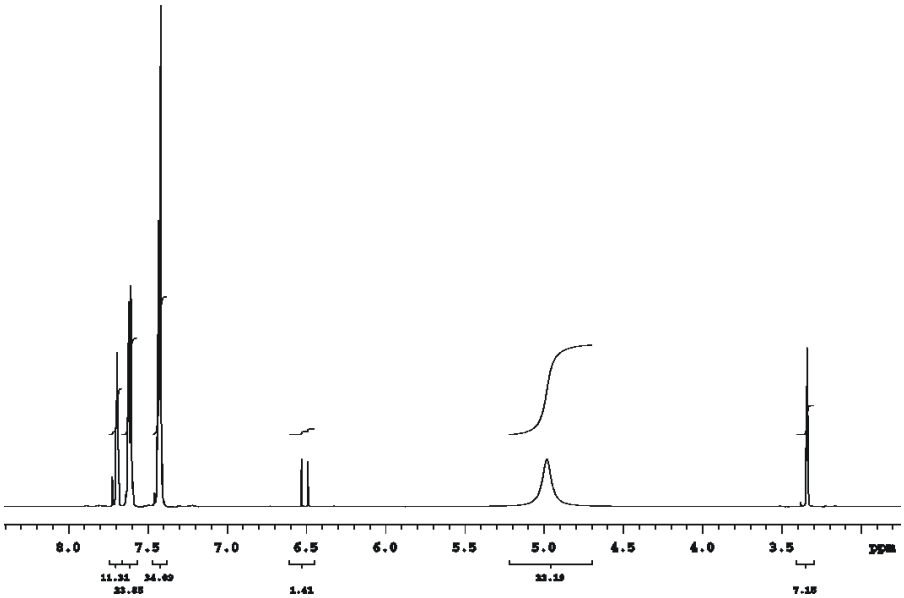


Fig. 3.4. ^{1}H NMR spectrum (CD_3OD, 400 MHz) of *E*-cinnamic-2-d_1 acid.

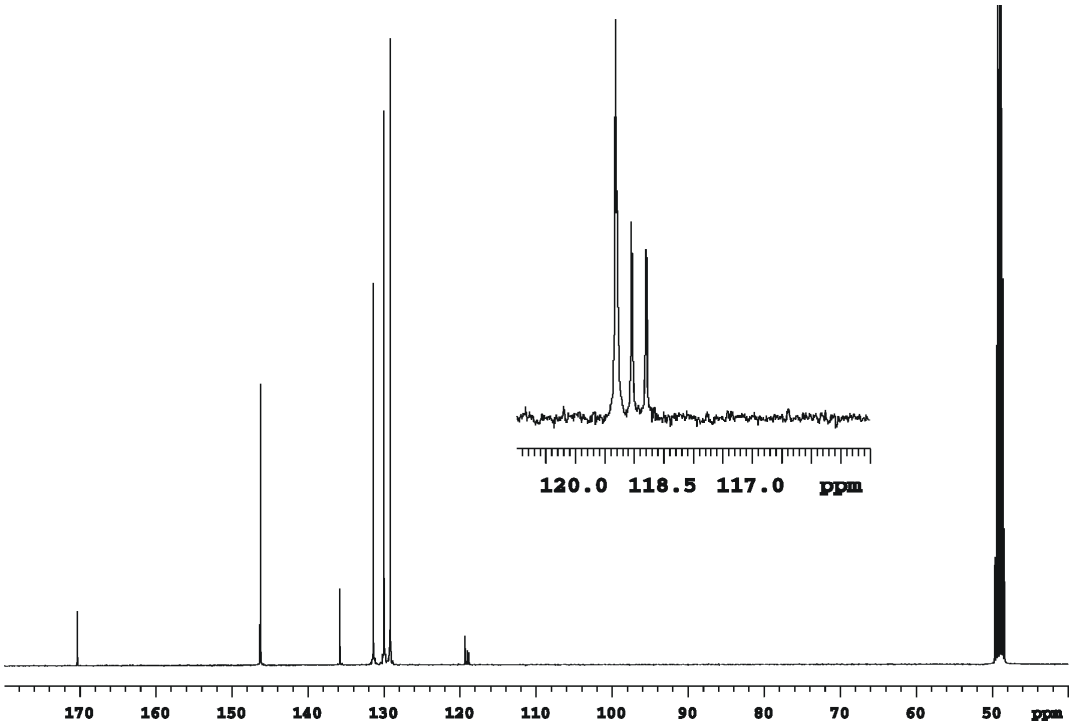


Fig. 3.5. ^{13}C NMR spectrum (CD_3OD, 100 MHz) of *E*-cinnamic-2-d_1 acid.

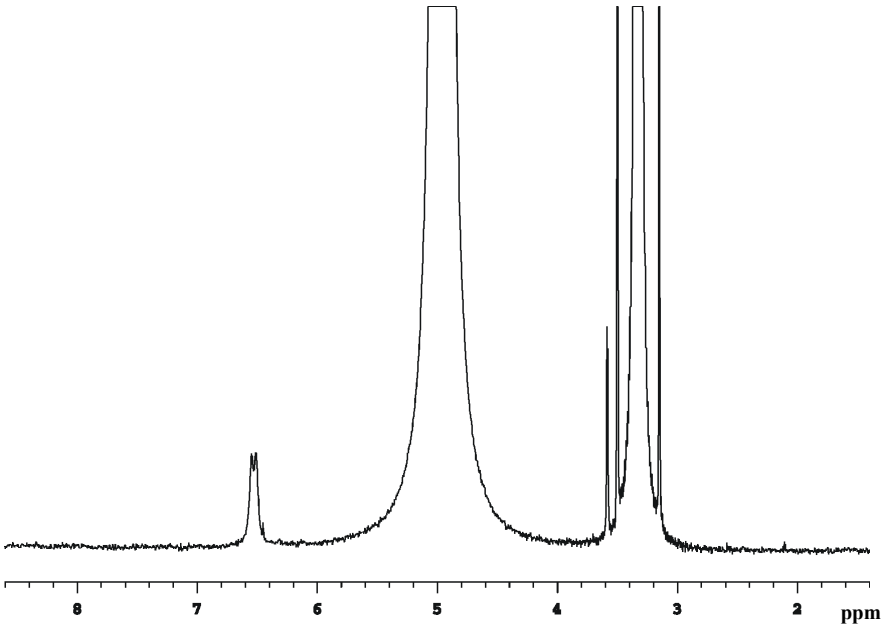


Fig. 3.6. ^{2}H NMR spectrum (CD_3OD, 61.41 MHz) of *E*-cinnamic-2-d_1 acid.

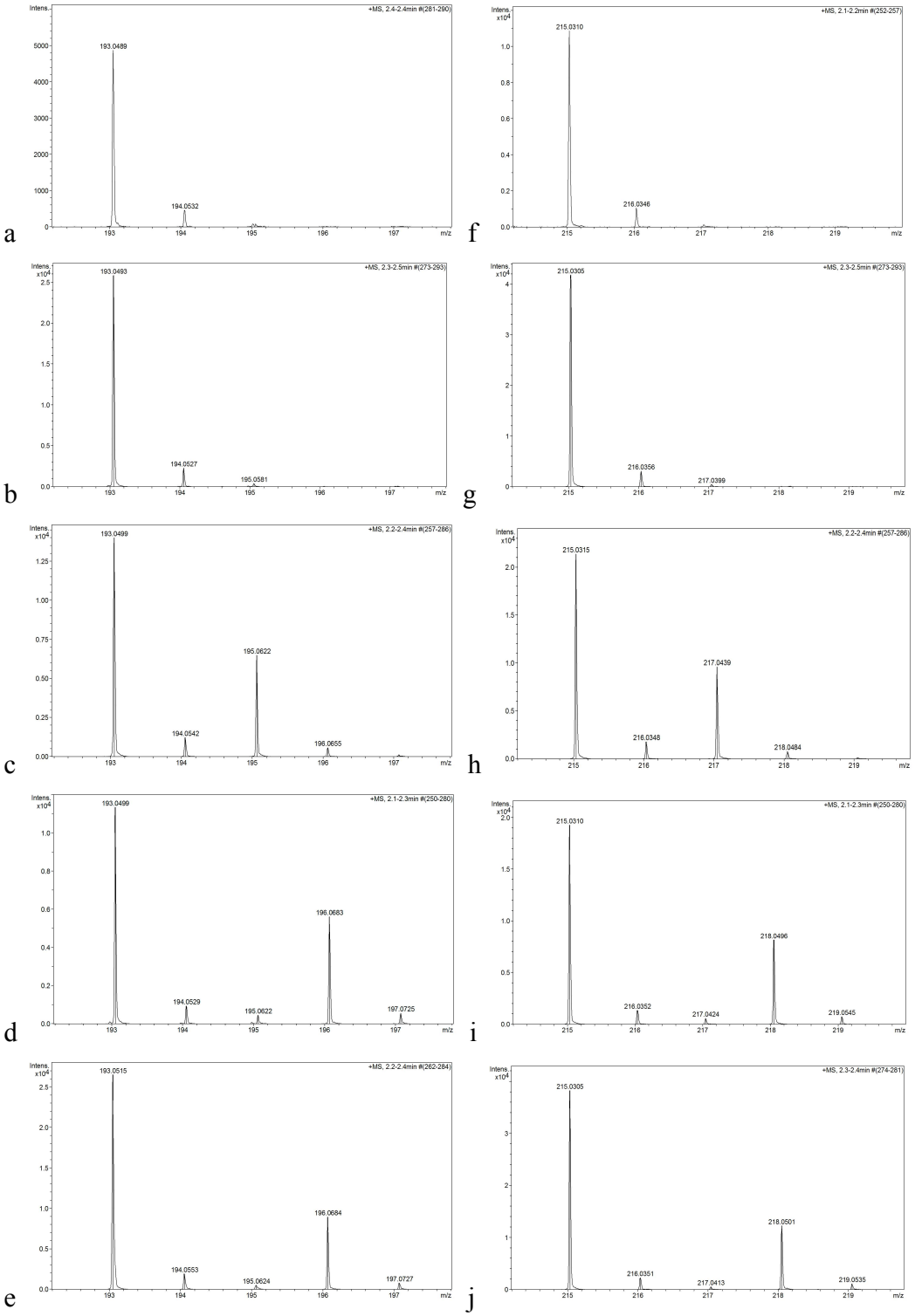


Fig. 3.7. HRMS of scopoletin products (M + ^{1}H) (a-e) and (M + ^{23}Na) (f-j) after feeding experiments with: (a) control, (b) *E*-cinnamate-2-d_1, (c) *E*-cinnamate-2',3',4',5',6'-d_5, (d) *E*-cinnamate-3,2',3',4',5',6'-d_6, (e) *E*-cinnamate-2,3,2',3',4',5',6'-d_7 acid and the scopoletin.

HR ESI MS spectra (Fig.3. 7) of scopoletin isolated from feeding experiments with *E*-cinnamic-2,3,2',3',4',5',6'-d_7 acid and *E*-cinnamic-3,2',3',4',5',6'-d_6 acid (Table 3.2) shows the presence of the natural peaks of scopoletin in addition to peaks of scopoletin-d_3. That there were no MS peaks corresponding to scopoletin-d_4 confirmed the preliminary result from the feeding experiment of *E*-cinnamic-d_7 acid. This loss of one deuterium atom more than expected during the biosynthesis requires a reconsideration of the mechanisms along the pathway. In order to determine which additional deuterium atom had been lost, a further feeding experiment was carried out using *E*-cinnamic-3,2',3',4',5',6'-d_6 acid, i.e. with ^{1}H isotope at C-2. This afforded the same two products as judged by both HPLC and HRMS. Feeding with *E*-cinnamic-d_6 acid did not result in an MS peak at 195 (scopoletin-d_2, $C_{10}H_7D_2O_4$ $[M + H]^+$) which was to be expected if the deuterium loss was from any position other than C-2 (Fig. 3.8).

In order to confirm this loss of proton or deuterium from C-2, we decided to feed with the deuterium isotope located at C-2. When we fed the cassava cubes with our synthetic *E*-cinnamic-2-d_1 acid (88% enriched) this gave HPLC peaks and corresponding HR MS data entirely consistent with unlabelled scopolin and scopoletin. Finally, in this series, feeding with *E*-cinnamic-2',3',4',5',6'-d_5 acid gave the same HPLC peaks, but HR MS data consistent with scopolin-d_2 and scopoletin-d_2 and therefore, apart from the 3 sites of oxygenation (2', 4', 5'), there was no further loss of deuterium from the aromatic ring. We therefore compared the percentage incorporation between the labelled and unlabelled hydroxycoumarins from these four feeding experiments (Tables 3.1and 3.2).

The HRMS results of the $[M + Na]^+$ ions (Fig. 3.7 f-j) agreed with the $[M + H]^+$ data (Fig. 3.7 a-e, Tables 3.1 and 3.2), typically: 215.0310 ($C_{10}H_8O_4Na$ requires 215.0315) from feeding with no label (and also from *E*-cinnamic-2,-d_1 acid), 218.0501 ($C_{10}H_5D_3O_4Na$ requires 218.0503) from *E*-cinnamic-2,3,2',3',4',5',6'-d_7 acid (and also from *E*-cinnamic-3,2',3',4',5',6'-d_6 acid), and 217.0439 ($C_{10}H_6D_2O_4Na$ requires 217.0440) from *E*-cinnamic-2',3',4',5',6'-d_5. Also, feeding with *E*-cinnamic-d_7 acid gave an HRMS peak of $[2M + Na]^+$, $C_{20}H_{16}O_8Na$ requires 407.0737, found 407.0729 and it's isotopic labelled peak $[2M + Na]^+$, $C_{20}H_{13}D_3O_8Na$ requires 410.0921, found 410.0926 $[M + Na]^+$, which confirm the result of formation of scopoletin-d_3.

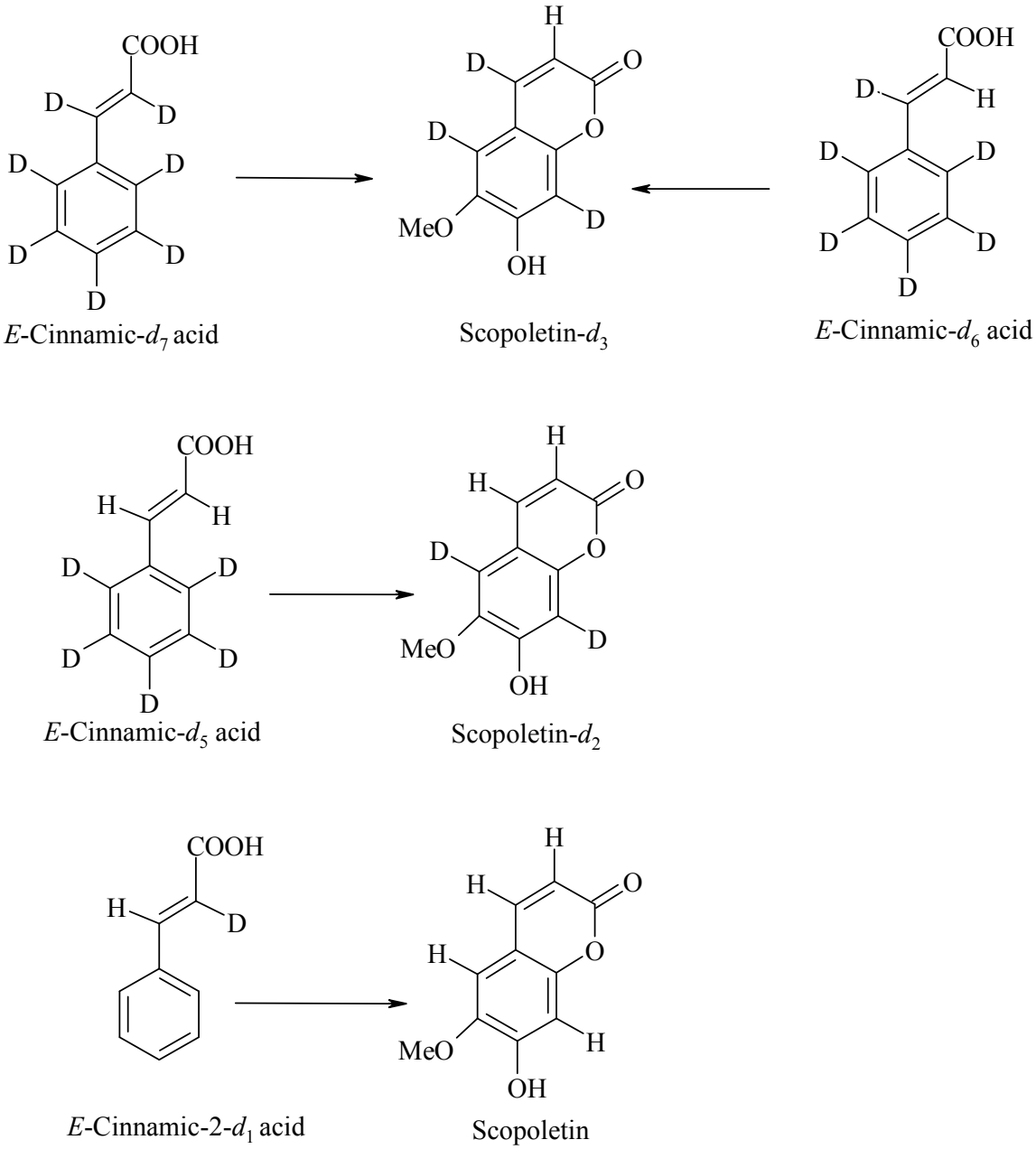


Fig. 3.8. Feeding experiments with *E*-cinnamic-2,3,2',3',4',5',6'-d_7 acid, *E*-cinnamic-3,2',3',4',5',6'-d_6 acid, *E*-cinnamic-2',3',4',5',6'-d_5 acid and *E*-cinnamic-2-d_1 and the isotopic label patterns found in product scopoletin.

Table 3.1

HR MS data of the HPLC peak at 7.1 min

Feeding cassava cv MCOL 22 with	***m/z* found of scopolin**	**Isotope peak monitored**	***m/z* found of scopolin-$d_{1\text{-}3}$ (calcd.)**	**% of isotope peak/*m/z* 355**
None (natural abundance ^{13}C)	355.1033	$C_{15}{}^{13}CH_{19}O_9$	356.1075 (356.1057)	13.9
Cinnamic-d_1 acid	355.1038	$C_{16}H_{18}D_1O_9$	356.1075 (356.1086)	13.9
Cinnamic-d_5 acid	355.1059	$C_{16}H_{17}D_2O_9$	357.1186 (357.1149)	32.3
Cinnamic-d_6 acid	355.1108	$C_{16}H_{16}D_3O_9$	358.1245 (358.1212)	34.0
Cinnamic-d_7 acid	355.1148	$C_{16}H_{16}D_3O_9$	358.1214 (358.1212)	26.1

scopolin, $C_{16}H_{19}O_9$ $[M + H]^+$ (calcd. 355.1024)

Table 3.2

HR MS data of the HPLC peak at 24.7 min

Feeding cassava cv MCOL 22 with	***m/z* found of scopoletin**	**Isotope peak monitored**	***m/z* found of scopoletin-$d_{1\text{-}3}$, (calcd.)**	**% of isotope peak/*m/z* 193**
None (natural abundance)	193.0489	$C_9{}^{13}CH_9O_4$	194.0532 (194.0529)	7.5
Cinnamic-d_1 acid	193.0493	$C_{10}H_8D_1O_4$	194.0527 (194.0558)	8.3
Cinnamic-d_5 acid	193.0499	$C_{10}H_7D_2O_4$	195.0622 (195.0621)	31.8
Cinnamic-d_6 acid	193.0499	$C_{10}H_6D_3O_4$	196.0683 (196.0684)	33.1
Cinnamic-d_7 acid	193.0515	$C_{10}H_6D_3O_4$	196.0684 (196.0684)	25.0

scopoletin, $C_{10}H_9O_4$ $[M + H]^+$, (calcd. 193.0495)

HPLC separation of the peak at 7.1 min is consistent with scopolin and HR ESI MS for this peak in each experiment is summarized in Table 3.1. HPLC separation of the peak at 24.7 min is consistent with the authentic standard scopoletin and HR ESI MS for this peak in each experiment is also summarized in Table 3.2. The percentages of labelled scopoletin and scopolin were calculated and summarised (Tables 3.1 and 3.2). These results (Fig. 3.8) confirmed that the *E*-*Z*-isomerisation of the C=C double bond in the biosynthesis of scopoletin and scopolin in cassava roots during PPD involves the specific exchange of the hydrogen atom at position 2 of *E*-cinnamic acid and thus is likely to be enzyme catalysed. If this biosynthetic isomerisation step along the pathway to scopoletin was a non-enzymatic photochemical reaction, it would be expected to proceed through a diradical, and there would be no loss from C-2 of deuterium (or proton) (Fig. 3.9). Feeding experiments with *E*-cinnamic acid-d_7 in light or in dark conditions were therefore carried out to determine whether or not the enzymatic reaction required light. The percentage of scopoletin-d_3 recovered in light was 9.8% while in dark was 9.1%. As there is no significant difference in the amount of scopoletin-d_3 biosynthesised in light or in dark conditions, the enzymatic step is light independent.

The E-cinnamic acid isomerisation stage in different plants

In order to interpret this loss of deuterium at position 2 when *E*-cinnamic-d_7 acid was fed to cassava root under PPD, we undertook a detailed literature search on the *E*-*Z*-isomerisation step in coumarin biosynthesis. The mechanism of this step varies between genera and even between species. It has often been reported that the isomerisation may be induced by UV light *in vivo*, as has been demonstrated *in vitro* (Koenigs et al., 1993; Zheng et al., 1999). Photoisomerisation of the *E*-double bond, in *p*-coumarate, has been studied using yellow protein that ultimately mediates a phototactic response to blue light in certain purple bacteria *Ectothiorhodospira* (Ryan et al., 2002; Dugave and Demange, 2003). Feeding *Melilotus officinalis* shoots with *E*-*o*-coumaric acid-2-^{14}C in both dark and light conditions showed much more radioactivity in the coumarins isolated from shoots exposed to light than in the coumarins isolated from shoots kept in the dark. Edwards and Stoker therefore concluded that an isomerase enzyme is not involved in the isomerisation of *o*-coumaric acid in *M. officinalis* shoots (Edwards and Stoker, 1967).

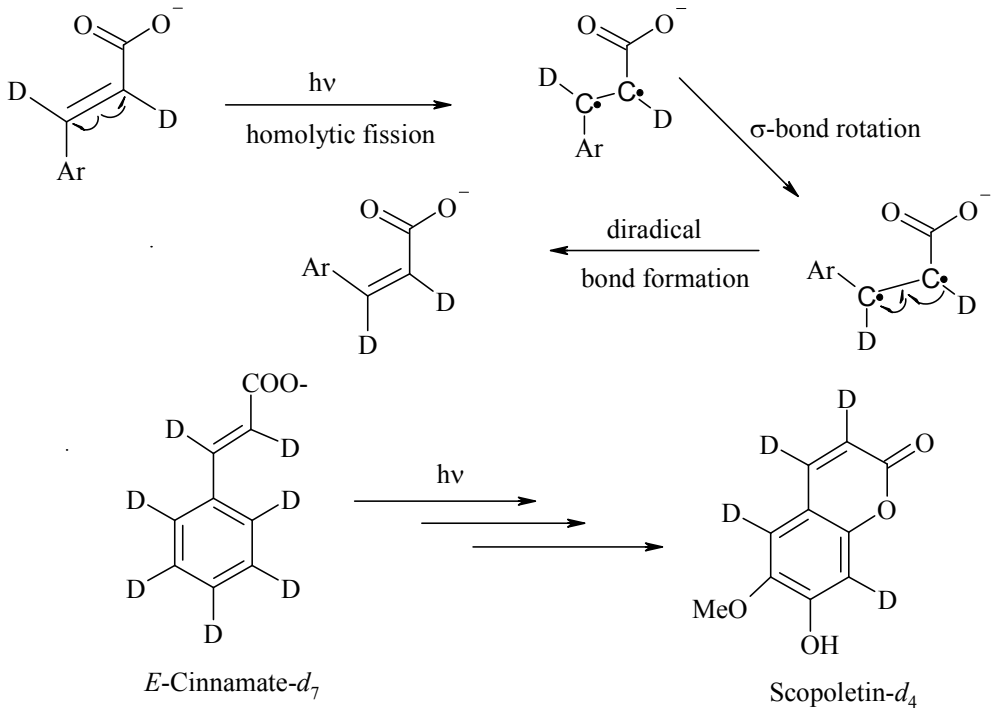


Fig. 3.9. Postulated photochemical isomerisation of *E*-cinnamic-d_7 acid.

In lavender (*Lavandula officinalis* and *L. spica*) (Edwards and Stoker, 1968), the biosynthesis of herniarin (7-methoxycoumarin) has been shown to be non-enzymatic, and the reaction is light catalysed. Therefore, based on the available evidence in 1967, Edwards and Stoker concluded that "it is probable therefore that the isomerisation step in the biosynthesis of all plant coumarins is entirely photochemical" (Edwards and Stoker, 1968). More recently, the isomerisation of cinnamic acid derivatives in barley and wheat was "directly attributed to the effect of light, and not apparently modulated by any enzymic reactions" (Turner et al., 1993), and in *Arabidopsis thaliana* "sunlight was able to isomerise both cinnamic acid isomers" (Wong et al., 2005).

Melilotus alba (sweet clover) plants fed with *E*-cinnamic acid-3-^{14}C in the dark produced ^{14}C-labelled coumarin. The amount of radioactivity in the coumarin was less than when plants were exposed to sunlight during the feeding experiments. *M. alba* leaves (homogenates) were also found to convert the 2'-β-glucoside of *o*-coumaric acid-2-^{14}C into coumarin in the dark. These results are indicative of the presence of an isomerase enzyme system in *M. alba* (Stoker, 1964). Moreover, there is evidence for

light-independent isomerisation in the vacuole (Strack, 1997). *E-Z*-Isomerisation of *o*-coumaric acid glucoside was studied in *M. alba* mespophyll cell protoplasts (Rataboul et al., 1985). Protoplasts were isolated and incubated in the presence of ^{14}C-labelled phenylalanine. After 4 h incubation in the light, ^{14}C-labelled *o*-coumaric acid glucoside was biosynthesised and transported to the vacuole. 60% of the synthesised *o*-coumaric acid glucoside was the *E*-isomer and 40% the *Z*-isomer. After longer incubation, 10.5 h in light, there was no significant change in the proportion of the two isomers. In contrast, if the protoplasts were incubated in light for 4 h, then transferred to the dark for 6.5 h, the amount of synthesised *o*-coumaric acid glucoside was decreased while the percentage of the *E*-isomer was 28% and the percentage of the *Z*-isomer was 72% which indicates that isomerisation increased. These experiments indicate that *E-Z* isomerisation of *o*-coumaric acid glucoside is not light dependent (Rataboul et al., 1985).

Conflicting literature views were revealed on the isomerisation step in coumarin biosynthesis which vary from one species to another and have not been reported previously for scopoletin and scopolin biosynthesis in cassava or in other plants. In our study on cassava roots, if the isomerisation step in scopoletin and scopolin biosynthesis is photochemical, scopoletin-d_4 and scopolin-d_4 should be biosynthesised in cassava roots during PPD when fed *E*-cinnamic-d_7 acid (Fig. 3.9). However, if the isomerisation is enzymatic, scopoletin-d_4 and scopolin-d_4 or scopoletin-d_3 and scopolin-d_3 could be biosynthesised depending on the enzyme mechanism. This has not been discussed before in the biosynthesis of any coumarins either in cassava or in other plants.

Studies with maleylacetone *Z-E*-isomerase proposed that glutathione-s-transferase could catalyse the addition of glutathione (GSH) acting as cellular nucleophile in the enzyme-catalyzed *Z-E*-isomerisation. The proposed intermediate is a dienediol formed by sulfur attack at C-2 (Fig. 3.10) and this intermediate undergoes σ-bond rotation and then elimination (possibly even before protonation at C-3 can take place) (Seltzer and Lin, 1979; Dixon et al., 2000). The key step (Fig. 3.11) is via a regiospecific 1,4- or 1,6-Michael addition to the enone, not to the unsaturated carboxylate. Alternatively, a nucleophilic addition of water (Fig. 3.12) across the C=C double bond may occur as in the second step of the well known β-oxidation of fatty acids (Jin et al., 1992), followed by σ-bond rotation and elimination of water to effect the isomerisation.

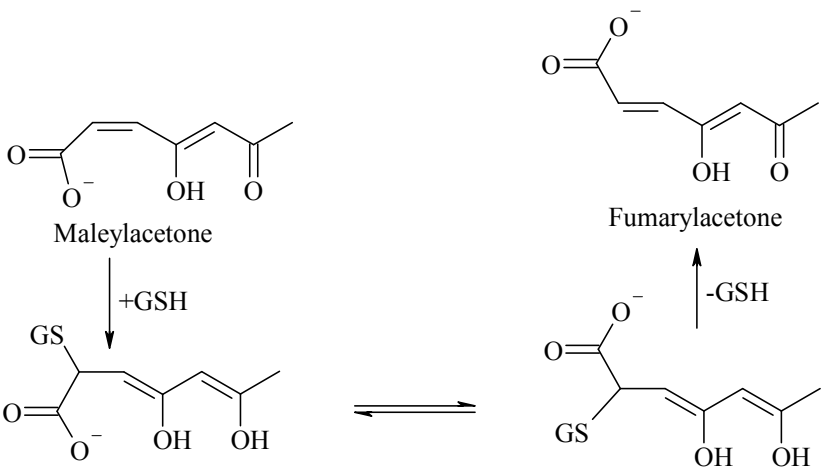


Fig. 3.10. Isomerisation of maleylacetone to fumarylacetone via formation of the GSH intermediate (regiospecific 1,4- or 1,6-Michael addition to the enone), (Seltzer and Lin, 1979; Dixon et al., 2000) σ-bond rotation and GSH elimination.

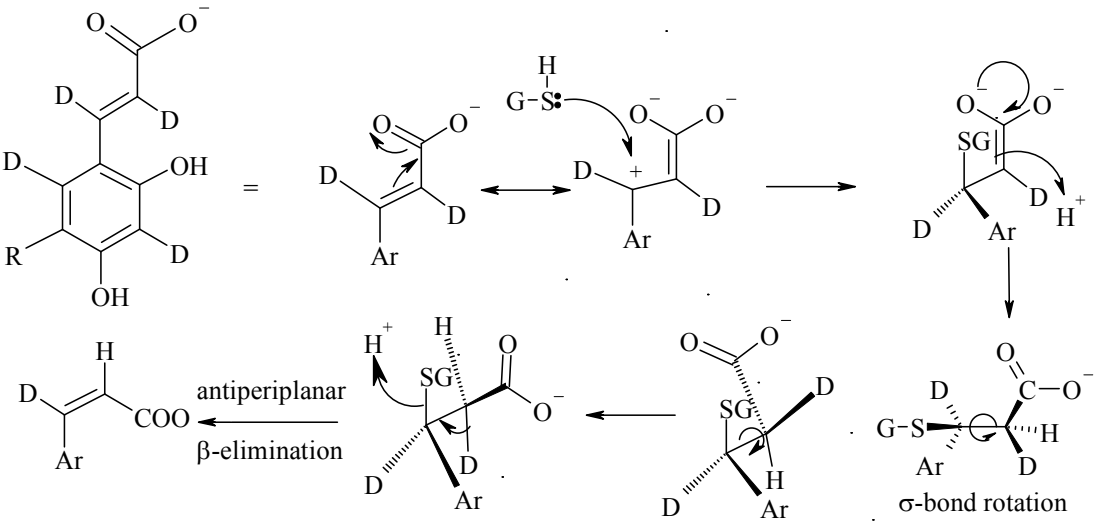


Fig. 3.11. Possible *E*-*Z* isomerisation mechanism of *E*-2,4-dihydroxycinnamate-d_7 and its substituents (R = D 2,4-dihydroxycinnamate; R = OH 2-hydroxycaffeate; R = OMe 6-hydroxyferulate) to their corresponding *Z*-2,4-dihydroxycinnamate-d_6 isomers via the formation of the GSH intermediate adducts shown by comparison with maleylacetone (Fig. 3.10).

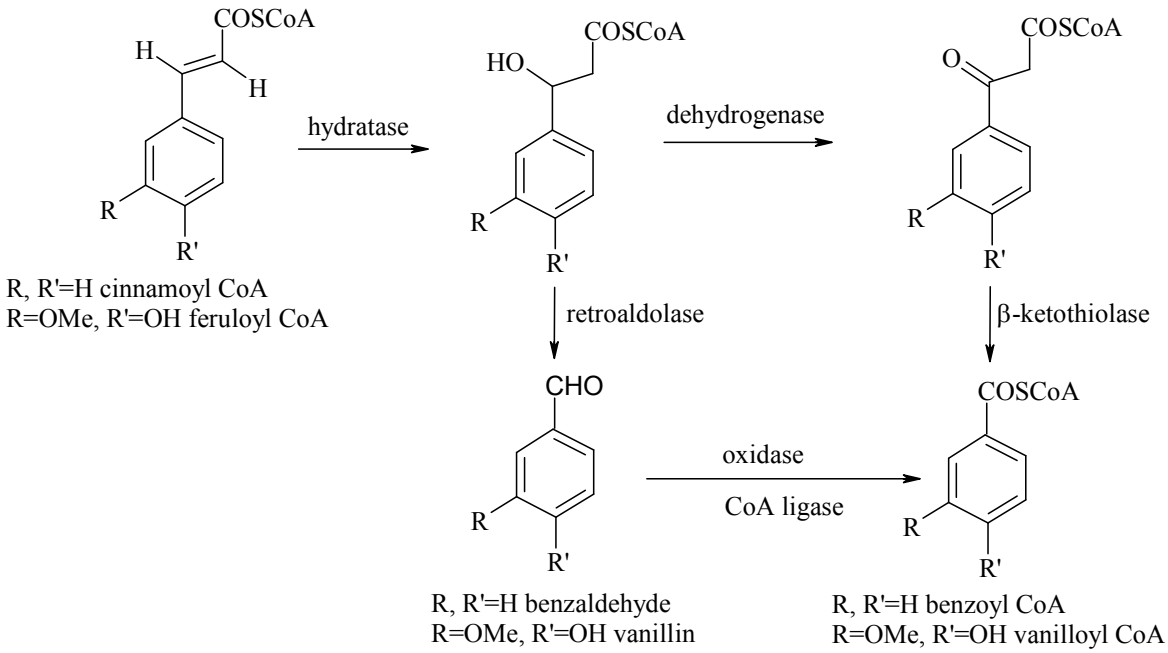


Fig. 3.12. Biosynthetic pathways of benzoic acid and vanillin (Gasson et al., 1998).

Water is similarly added during first step of benzaldehyde, vanillin, benzoic, salicylic and vanilloic acid biosynthesis from cinnamic acid derivatives. Enoyl-SCoA hydratase in plants and bacteria catalyses the addition of water across the double bond of the α,β-unsaturated thiol ester (e.g. cinnamoyl SCoA or feruloyl SCoA) (Gasson et al., 1998; Abd El-Mawla and Beerhues, 2002; Bahnson et al., 2002; Walton et al., 2003). The intermediate from such a nucleophilic addition at the β-position could undergo σ-bond rotation and then antiperiplanar- (β-) elimination leading to the formation of the geometrical isomer.

Since in the literature there was nothing about the mechanism of *E-Z* isomerisation step of cinnamic acid derivatives, this step could be similar to the isomerisation of maleylacetone; which involves the nucleophile addition of GSH and this intermediate undergoes σ-bond rotation and then antiperiplanar elimination. The elimination step involves loss of D at position 2. Alternatively, it could be through nucleophilic addition of water, σ-bond rotation and then antiperiplanar elimination. We therefore conclude that as scopoletin-d_3 and scopolin-d_3 (not d_4) were biosynthesised when *E*-cinnamic-d_7 acid was fed, and as we have established that the isomerisation stage in their biosynthesis in cassava roots during PPD is not photochemical, rather it is enzymatic and is catalysed by an isomerase enzyme that is not dependent on light.

Chapter 4

Biosynthesis of hydroxycoumarins in cassava roots during PPD: different pathways and the *o*-hydroxylation step

4.1. Introduction

Despite the importance of coumarins in plant development and defence, and with their wide variety of pharmacological activities, key aspects of the biosynthesis of these secondary metabolites remain unresolved (Petersen et al., 1999; Bourgaud et al., 2006). All potential biosynthetic routes from cinnamic acid to hydroxycoumarins require an *E*-*Z* isomerisation of the double bond discussed in chapter 3. However, several hypothetical pathways from cinnamate to hydroxycoumarin have been proposed, and evidence from other species suggests different routes may occur in different species. Thus we exploited the observation that the accumulation of scopoletin and its glucoside scopolin increases in cassava roots (*Manihot esculenta* Crantz Family Euphorbiaceae) during PPD (Buschmann et al., 2000b; Reilly et al., 2003) to test alternative pathways for the biosynthesis of these hydroxycoumarins.

Three hypothetical pathways for the biosynthesis of scopoletin (Kai et al., 2006) via: 2',4'-dihydroxycinnamate (Fig. 4.1), 3',4'-dihydroxycinnamate (caffeate) (Fig. 4.2), or 4'-hydroxy-3'-methoxycinnamate (ferulate) (Fig. 4.3) have been proposed from studies in various plant species. The observation that esculetin and its glucoside esculin accumulate in deteriorated cassava in lesser amounts than scopoletin and scopolin is consistent with a route via 2',4'-dihydroxycinnamate or caffeate to esculetin. In *Agathosma puberula*, ferulic, sinapic and caffeic acids were poorly utilised than *p*-coumaric acid in the biosynthesis of puberulin (6,8-dimethoxy-7-prenyloxocoumarin). Suggesting a pathway biosynthesised from umbelliferone via esculetin and then scopoletin (Brown et al., 1984; Brown et al., 1988). This implies hydroxylation and methylation after lactonisation. In addition, in *Cichorium intybus* (chicory) it was confirmed that umbelliferone was converted into esculetin which provided further support for the theory that polyoxygenated coumarins, in general, elaborated by additional oxygenation of umbelliferone (Brown, 1985).

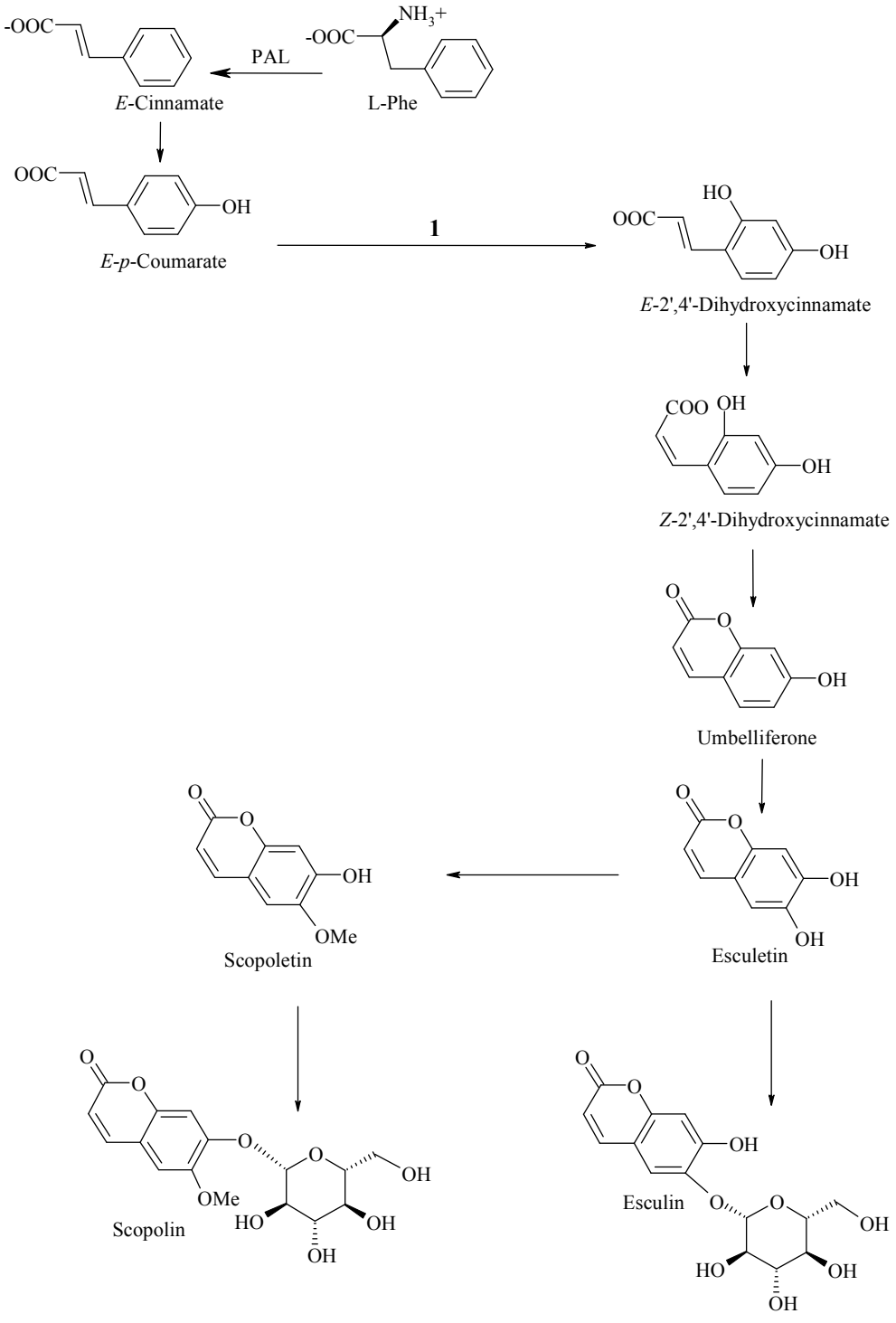


Fig. 4.1. Proposed metabolic pathway in the biosynthesis of scopoletin via 2',4'-dihydroxycinnamate (for clarity the carboxylate anion has been omitted after *E*-cinnamate).

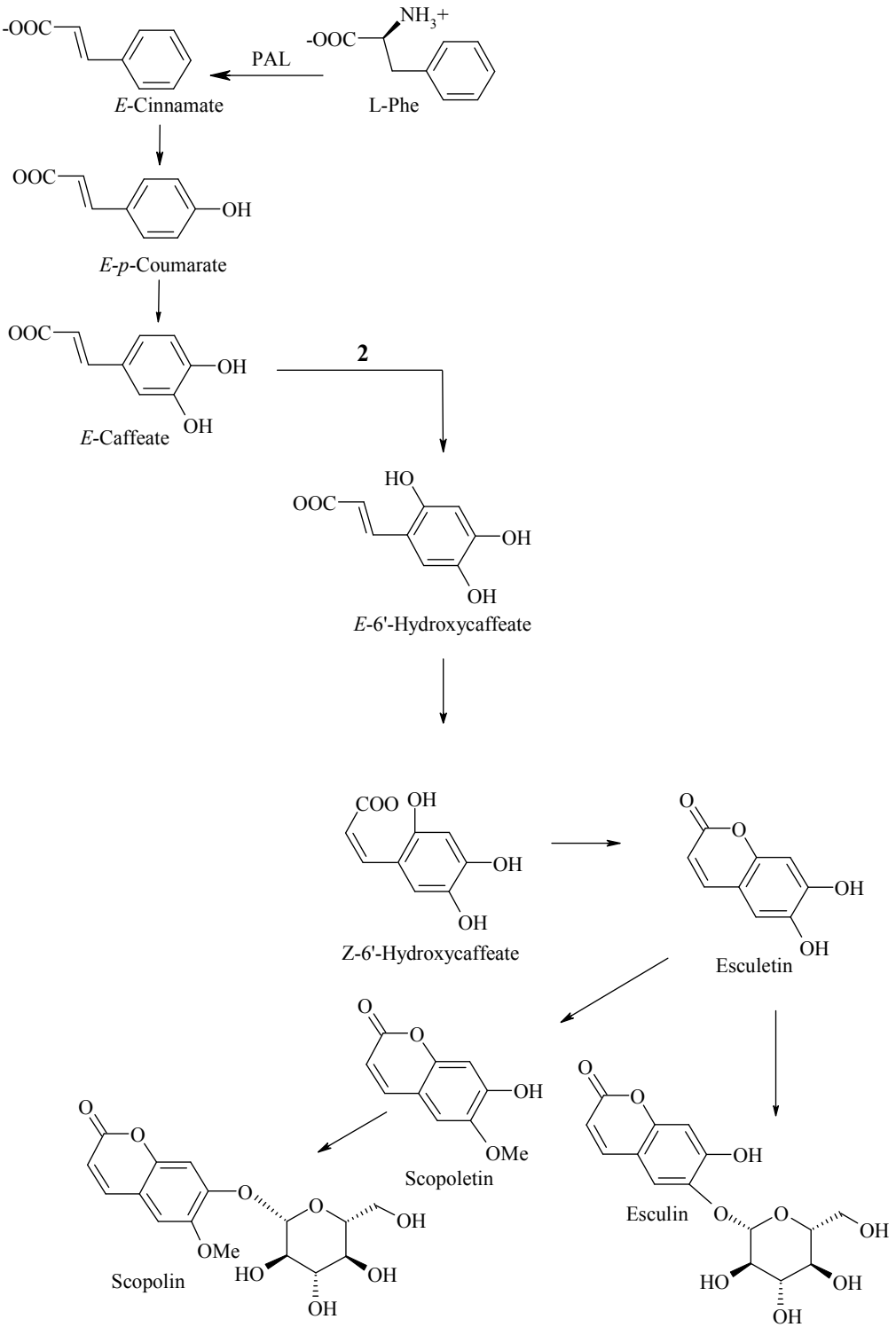


Fig. 4.2. Proposed metabolic pathway in the biosynthesis of scopoletin via caffeic acid.

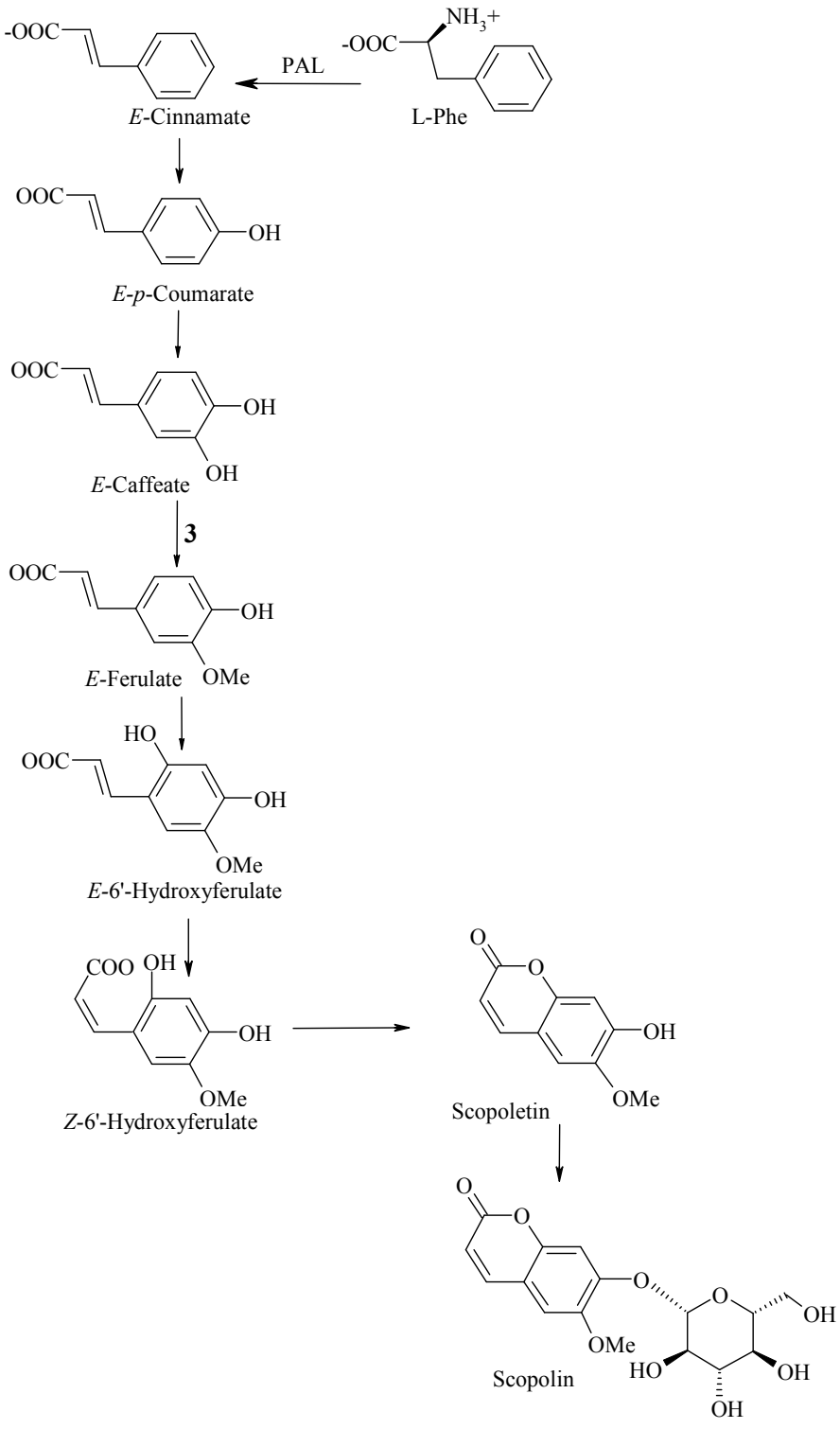


Fig. 4.3. Proposed metabolic pathway in the biosynthesis of scopoletin via ferulic acid.

However, from studies in *Arabidopsis thaliana* the content of scopoletin and scopolin in the mutant roots (T-DNA inserted within the gene encoding the cytochrome P450, CYP98A3, which catalyzes 3'-hydroxylation of *p*-coumarate) greatly decreased to similar to 3% of that in the wild-type roots which suggested that scopoletin and scopolin biosynthesis is strongly dependent on the 3'-hydroxylation of *p*-coumarate units catalyzed by CYP98A3 (Kai et al., 2006), this indicates that scopoletin is biosynthesised in *A. thaliana* from either caffeic (Fig. 4.2) or ferulic acid (Fig. 4.3), and recently feruloyl-CoA has been proposed as a key precursor in scopoletin biosynthesis in this plant (Kai et al., 2008). Also feeding *Nicotiana tabacum* (tobacco) with ^{14}C-labelled ferulate led to this being actively metabolised by the leaf tissues (Runeckles, 1963) or tobacco tissue cultures (Fritig et al., 1970) into scopoletin and its glucoside scopolin which suggests methylation prior to lactonisation.

In conclusion, from the previous literature the biosynthesis of scopoletin and scopolin in cassava roots during PPD could be via *p*-coumaric acid (pathway 1, Fig. 4.4) which is *o*-hydroxylated to 2',4'-dihydroxycinnamic acid then isomersed and lactonised to umbelliferone then hydroxylated to esculetin and methylated to scopoletin and or via caffeic acid (pathway 2, Fig. 4.4) which is *o*-hydroxylated to 6'-hydroxycaffeic acid then isomerised and lactonised to esculetin then methylated into scopoletin and or via ferulic acid (pathway 3, Fig. 4.4) which is *o*-hydroxylated to 6'-hydroxyferulate then isomerised and lactonised.

We have previously shown that when chopped cassava roots were fed with deuteriated *E*-cinnamic acid, and PPD allowed to occur, scopoletin containing a certain percentage of labelled scopoletin was produced. In order to investigate the pathway of scopoletin biosynthesis in cassava roots during PPD, direct and competition feeding experiments with other potential intermediates were carried out. Cassava roots under PPD were fed with stable isotopically labelled *p*-coumaric (4'-hydroxycinnamic) acid, caffeic acid, ferulic acid, umbelliferone and esculetin. In addition, competition feeding experiments were carried out by feeding the root samples with deuteriated *E*-cinnamic -(d_5 and d_7) acid alone as a control and with deuteriated *E*-cinnamic-(d_5 and d_7) acids mixed with putative unlabelled intermediates along the scopoletin biosynthetic pathway which could compete with deuteriated *E*-cinnamic acid.

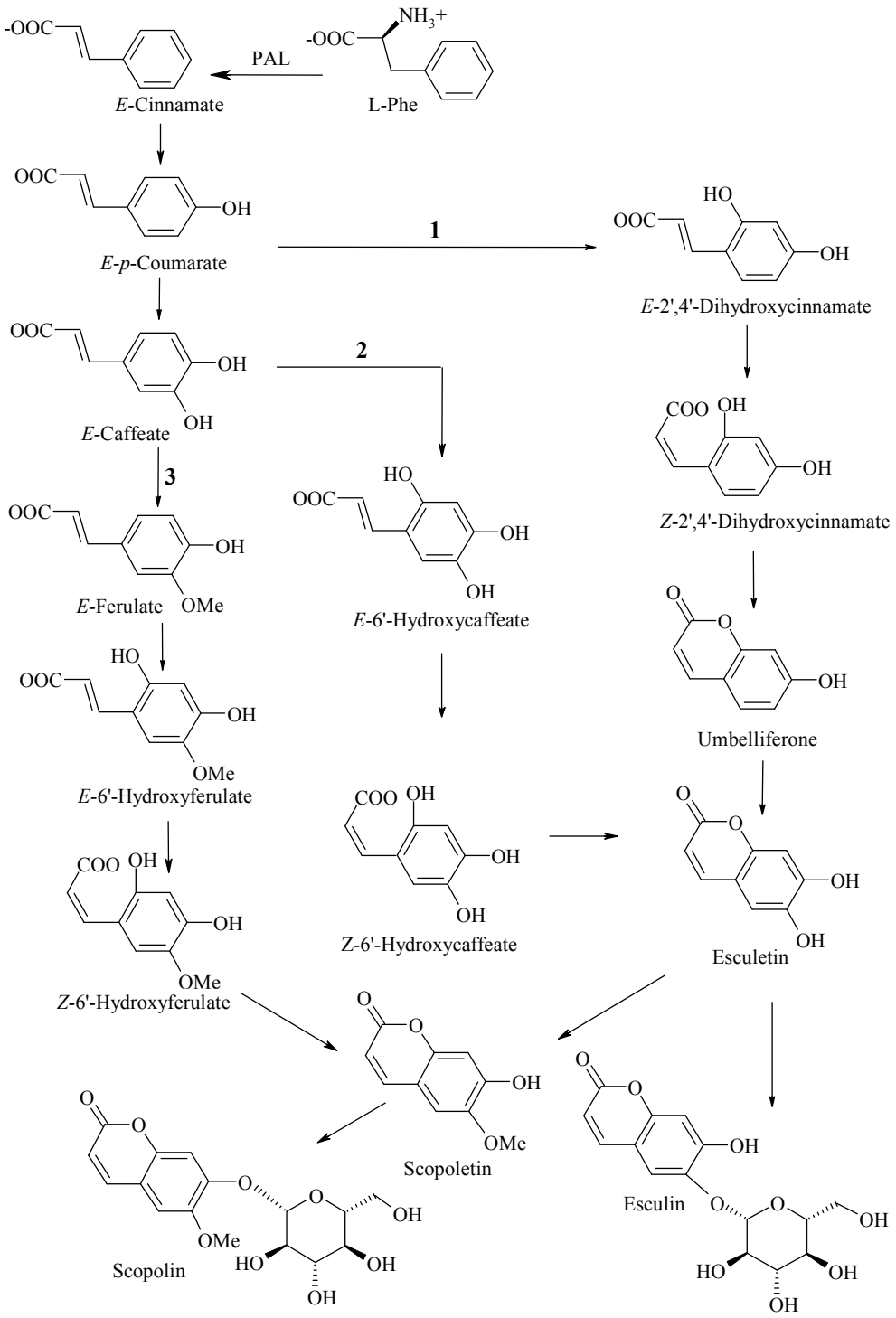


Fig. 4.4. Three proposed metabolic pathways in the biosynthesis of scopoletin.

Two routes have been proposed for the key benzo-2-pyrone cyclisation reaction of 4'-hydroxycinnamic acid (Fig. 4.5) (Petersen et al., 1999). Oxidative cyclisation via a spirodienone intermediate would yield umbelliferone in which the lactone ether oxygen is derived from the carboxyl group. The spirolactone-dienone intermediate pathway has been previously established in *Streptomyces niveus* for novobiocin biosynthesis in elegant work by Kenner and co-workers (Bunton et al., 1963) and also proposed from UV studies in cultures of the plant *Ammi majus* L. (Apiaceae, Bishop's flower, large bullwort) (Matern, 1991) following work by (Grisebach and Ollis, 1961). Alternatively the 2'-hydroxylation of 4'-hydroxycinnamic acid may be followed by *E*- to *Z*-isomerisation and lactonisation, when the lactone ether oxygen will be derived from the 2'-hydroxy group and not from the carboxylic acid. Although *o*-hydroxylation of cinnamic (*p*-coumaric, caffeic or ferulic) acids is of pivotal importance for the biosynthesis of coumarins, it remains a missing link in phenylpropanoid biosynthesis (Bourgaud et al., 2006).

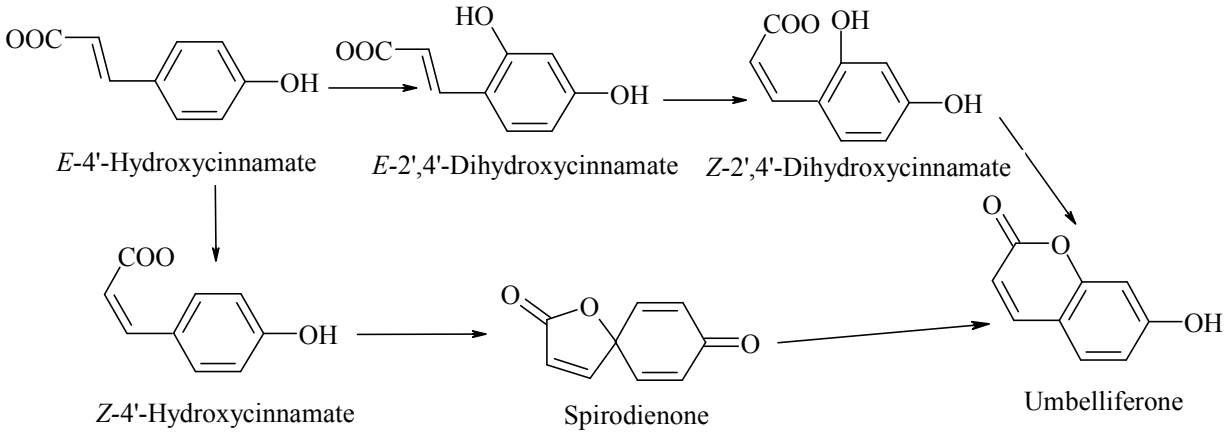


Fig. 4.5. Proposed routes for the conversion of 4'-hydroxycinnamate into umbelliferone

Support for 2'-hydroxylation came from observation of an NIH-shift in some studies. Aromatic hydroxylation may involve arene oxide intermediate formation. When the epoxide opens up the deuterium attached to the position which become hydroxylated can migrate to the adjacent carbon in the ring (Fig. 4.6). This hydrogen shift is known as an NIH-shift (from the National Institute of Health, Bethesda, USA where the phenomenon was first observed) is an excellent tool in the prediction of isotopic patterns in *o*-hydroxylation in phenylpropanoids (Schmidt et al., 2006). Labelled cinnamic acid (*o*-^{3}H, ring-1-^{14}C) was fed to *Melilotus alba* shoots and the

relative retention of tritium was 78% upon conversion into coumarin (Ellis and Amrhein, 1971). In *Gaultheria procumbens* leaves 92% relative retention of tritium was observed upon conversion into *o*-coumaric acid. These results imply the occurrence of a 2'-hydroxylation step in these species (Ellis and Amrhein, 1971). Recently, in *A. thaliana* it was shown that a Fe(II)- and 2-oxoglutarate-dependent dioxygenase (rather than a P450 enzyme) catalyses the *o*-hydroxylation of feruloyl-CoA in scopoletin biosynthesis (Kai et al., 2008).

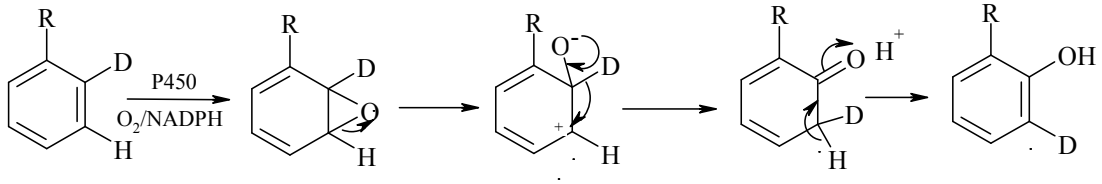


Fig. 4.6. The mechanism of NIH shift during aromatic *o*-hydroxylation

It is notable that the biosynthesis of salicylic acid and the origin of its *o*-hydroxyl group is still also a matter of debate with differences between species (Shah, 2003). In *N. benthamiana*, it was determined that salicylic acid is biosynthesised from isochorismate (Catinot et al., 2008), while in PAL-suppressed *N. tabacum* the level of salicylic acid was decreased (Pallas et al., 1996) suggesting that salicylic acid was synthesised from phenylalanine (Watts et al., 2004) through benzoic acid (Hertweck et al., 2001) which is then *o*-hydroxylated to salicylic acid. Following the ^{18}O-labelling of salicylic acid, it was proposed that benzoic acid *o*-hydroxylase is an oxygenase that specifically hydroxylates the *o*-position of benzoic acid. This hydroxylase enzyme was purified and shown to belong to a novel class of soluble, high molecular weight cytochrome P450 enzymes (Leon et al., 1995).

In order to investigate the origin of the lactone ether oxygen in coumarins biosynthesised in cassava roots during PPD, we have designed and carried out feeding experiments with $C^{18}O_2$-enriched *E*-cinnamic and ferulic acids by spraying an aqueous solution of each labelled cinnamic acid on cubes of freshly harvested cassava roots. Separately, in an atmosphere rich in $^{18}O_2$, the scopoletin and scopolin labelling patterns resulting from deuteriated cinnamic-d_7 acid feeding were obtained and compared with those from untreated roots kept in the same atmosphere.

4.2. Experimental

Plant material

Root tubers of different cultivars (MCOL 22, MNGA 19, MNGA 2) were harvested from cassava plants growing in the tropical glass house at the University of Bath as given above in Chapter 2.

General methods

Chemicals were obtained routinely from Sigma-Aldrich Chemical Co. Ltd, UK, except *E*-cinnamic-2',3',4',5',6'-d_5 acid which was obtained from CDN Isotopes, Canada and $H_2{}^{18}O$ 95 atom % ^{18}O from CK GAS Products Ltd. The HPLC instrument was a Jasco PU-980 pump, monitored at 360 nm with a Jasco UV-975 detector, using 16% acetonitrile : 84% aq. formic acid (0.1%), flow rate 4 ml/min at 20 °C. HPLC data were recorded on a Goerz Metrawatt Servogor 120 recorder, HPLC columns, purchased from Phenomenex Inc.: Phenomenex Gemini 10 μ C18 110A 250 x 10 mm with guard column Phenomenex Gemini 5 μ C18 10 x 10 mm. Samples were injected using a 100 μl loop. HR ESI MS was carried out on a Bruker microTOF mass spectrometer in the Department of Pharmacy and Pharmacology, University of Bath or on a Micromass Quattro II in the EPSRC National Mass Spectrometry Service Centre, University of Wales, Swansea. NMR spectra were obtained on a Varian Mercury Spectrometer at 400 MHz (^{1}H) and 100 MHz (^{13}C) in CD_3OD, all chemical shifts are reported in parts per million (ppm) relative to internal tetramethylsilane, and coupling constants (*J*) are absolute values in Hz.

General feeding procedure

Cassava roots (typically 1 kg) were peeled, then cut into approximately 1 cm^3 cubes, divided into groups (typically 100 g) and treated as given above in Chapter 3.

Synthesis of p-coumaric-2-^{13}C acid, caffeic-2-^{13}C acid and ferulic-2-^{13}C acid

p-Hydroxybenzaldehyde, 3,4-dihydroxybenzaldehyde or vanillin (4-hydroxy-3-methoxybenzaldehyde) (1.2 mmol) and malonic-2-^{13}C acid (99 atom % ^{13}C) (276 mg,

2.63 mmol) were dissolved in pyridine (660 µl). Piperidine (13 µl, 130 nmol (0.00011 eq.) was added, the reaction was heated to 70 ^{o}C (oil bath) for 24 h. Water (7.5 ml), and then concentrated HCl (0.4 ml), were added dropwise until a precipitate appeared which was collected and recrystallised. The product purity was checked by TLC with n-hexane-ethyl acetate-acetic acid (1:1:0.01, v/v/v).

p-Coumaric-2-^{13}C acid, white crystals (H_2O) m.p. 208-210 ºC (61% relative to *p*-hydroxybenzaldehyde), TLC $R_f = 0.36$. $C_8{}^{13}CH_7O_3$ requires 164.0434, HR ESI MS gave *m/z* 164.0434 $[M - H]^-$. ^{1}H NMR and ^{13}C NMR data see: Tables 4.1 and 4.2.

Caffeic-2-^{13}C acid, yellowish-white crystals (H_2O) m.p. 217-220 ºC (64% relative to 3,4-dihydroxybenzaldehyde), TLC $R_f = 0.22$. $C_8{}^{13}CH_7O_4$ requires 180.0383, HR ESI MS gave *m/z* 180.0381 $[M - H]^-$. ^{1}H NMR and ^{13}C NMR data, see: Tables 4.1 and 4.2.

Ferulic-2-^{13}C acid, white crystals (H_2O) m.p. 168-171 ºC (64% relative to vanillin), TLC $R_f = 0.35$. A single crystal X-ray of the labelled product is shown (Fig. 4.8). $C_9{}^{13}CH_9O_4$ requires 194.0540, HR ESI MS gave *m/z* 194.0538 $[M - H]^-$. ^{1}H NMR and ^{13}C NMR data, see: Tables 4.1 and 4.2.

X-Ray crystallography

A single crystal of the compound ferulic -2-^{13}C acid was analyzed on a Nonius-kappaCCD diffractometer in the Department of Chemistry, University of Bath, using graphite monochromated Mo-K_a radiation of wavelength 0.71073Å. The sample was maintained at 150 K throughout the data collection, using an Oxford Cryosystems cryostat. Details of the data collection and refinement are given in Appendix 1.

^{18}O-Enrichment of umbelliferone and esculetin

Following the literature procedure, (Chandrasekaran et al., 1984; Kollenz et al., 1991) umbelliferone (20 mg, 0.124 mmol) or esculetin (20 mg, 0.112 mmol), $H_2{}^{18}O$ (0.5 ml) (95 atom % ^{18}O), conc. HCl (10 µl) in acetonitrile (1 ml) were maintained at 70 ºC in a sealed tube for 3 h, 4 d, and 7 d. ^{18}O-Enriched umbelliferone, $C_9H_7O_2{}^{18}O$ requires 165.0432, HR ESI MS found 165.0430 $[M + H]^+$, $C_9H_7O_3$ requires 163.0390, found 163.0430 $[M + H]^+$. ^{18}O-Enriched esculetin, $C_9H_7O_3{}^{18}O$ requires 181.0381, HR MS found 181.0373 $[M + H]^+$ and $C_9H_7O_4$ requires 179.0339, found 179.0328 $[M + H]^+$.

Percentage of ^{18}O-enrichment of umbelliferone (% of *m/z* 165/163) was 8% (after 3 h), 61% (after 4 d), and 82% (after 7 d). Percentage of ^{18}O-enrichment of esculetin was 82% (after 7 d).

^{18}O-Enrichment of E-cinnamic acid and ferulic acid

E-Cinnamic acid (20 mg, 0.135 mmol) or ferulic acid (20 mg, 0.103 mmol), $H_2{}^{18}O$ (0.5 ml) (95 atom % ^{18}O) (from CK GAS Products Ltd) and conc. HCl (10 µl) in acetonitrile (1 ml) were maintained at 70 ºC in a sealed tube in for 3 h, 1, 2 and 4 d. ^{18}O-Enriched *E*-cinnamic acid after 4d, $C_9H_9O^{18}O$ requires 151.0640, HR ESI MS found 151.0662 $[M + H]^+$ and $C_9H_9{}^{18}O_2$ requires 153.0682, found 153.0687 $[M + H]^+$. The percentage of ^{18}O-enrichment *E*-cinnamic acid after each period is listed in Table 4.3. ^{18}O-Enriched *E*-ferulic acid after 4 d, $C_{10}H_{11}O_4$ requires 195.0652, HR MS found 195.0661 $[M + H]^+$ 2%, $C_{10}H_{11}O_3{}^{18}O$ requires 197.0694, found 197.0705 $[M + H]^+$ 20%, and $C_{10}H_{11}O_2{}^{18}O_2$ requires 199.0737, found 199.0743 $[M + H]^+$ 78%.

Feeding experiments with p-coumaric-2-^{13}C acid, caffeic-2-^{13}C acid and ferulic-2-^{13}C acid

Using the General Feeding procedure, cassava roots (0.8 kg, cv MCOL 22) were peeled (0.64 kg) and divided into four groups, one as a control (40 g, the roots left to deteriorate for 3 d and 4 d without feeding with any intermediates) and three equal groups (85 g), fed with *p*-coumaric-2-^{13}C acid, caffeic-2-^{13}C acid, and ferulic-2-^{13}C acid (20 mg of each acid) dissolved in aq. 4% Na_2CO_3 (2 ml). A representative sample of the combined EtOH extract (2 g) was then purified by HPLC. HR MS data of the HPLC peaks at 7.1 and 24.7 min of scopolin and scopoletin are listed in Tables 4.4 and 4.5.

Feeding experiments with ^{18}O-labelled esculetin and umbelliferone and ferulic-2-^{13}C acid

Using the General Feeding procedure, cassava roots (0.69 kg, cv MNGA 19) were peeled (0.49 kg) and divided into three equal groups (85 g). One was fed with ^{18}O-carbonyl labelled esculetin (20 mg), one with ^{18}O-carbonyl labelled umbelliferone (20 mg), and the other with ferulic-2-^{13}C acid (20 mg), each sample dissolved in DMSO (1 ml) then diluted with water (1 ml). A representative sample of the combined EtOH extract (2 g except group one 0.7 g) was then purified by HPLC. HR MS data of the

HPLC peaks at 6.0, 7.1 and 24.7 min of esculin, scopolin and scopoletin are listed in Tables 4.6, 4.7 and 4.8.

Competition feeding experiments between deuterium labelled E-cinnamic-d_7 acid and unlabelled intermediates

Using the General Feeding procedure, cassava roots (1.23 kg, cv MCOL 22) were peeled (1 kg) and divided into four equal groups, the first group was fed with *E*-cinnamic-d_7 acid (30 mg) dissolved in aq. 4% Na_2CO_3 (3 ml). The second, third and fourth groups were fed with *E*-cinnamic-d_7 acid (30 mg) dissolved in aq. 4% Na_2CO_3 (3 ml) then unlabelled caffeic, ferulic or *E*-2',4'-dihydroxycinnamic acids (30 mg each) prepared in the same way as the substrate was added. A representative sample of the combined EtOH extract (4 g) was then purified by HPLC. HR MS data of the HPLC peaks at 7.1 and 24.7 min of scopolin and scopoletin are listed in Tables 4.9 and 4.10.

Competition feeding experiments between deuterium labelled E-cinnamic-2',3',4',5',6'-d_5 acid and unlabelled umbelliferone and esculetin

Using the General Feeding procedure, cassava roots (0.68 kg, cv MCOL 22) were peeled (0.46 kg) and divided into three equal groups (115 g), the first was fed with *E*-cinnamic-d_5 acid (20 mg) dissolved in 1 ml of DMSO then 1 ml of water was added. The second group was fed with *E*-cinnamic-d_5 acid (20 mg) and umbelliferone (20 mg) dissolved in DMSO (1 ml) then water (1 ml) was added. The third group was fed with *E*-cinnamic-d_5 acid (20 mg) and esculetin (20 mg) dissolved in 1 ml of DMSO then 1 ml of water was added. A representative sample of the combined EtOH extract (3.5 g) was then purified by HPLC. HR MS data of the HPLC peaks at 7.1 and 24.7 min of scopolin and scopoletin are listed in Tables 4.11 and 4.12.

Feeding experiments with E-$C^{18}O_2$-cinnamic acid and $C^{18}O_2$-ferulic acid

Using the General Feeding procedure, cassava roots (0.69 kg, cv MNGA 19) were peeled (0.49 kg). One group (85 g) was fed with $C^{18}O_2$-*E*-cinnamic acid (20 mg) and the other group (65 g) was fed with $C^{18}O_2$-ferulic acid (10 mg) dissolved in aq. 4% Na_2CO_3 (2 ml). A representative sample of the combined EtOH extract (2 g) was then

purified by HPLC. From the feeding experiments with $C^{18}O_2$-*E*-cinnamic acid, the HPLC peak at 7.1 min scopolin, $C_{16}H_{19}O_9$ requires 355.1024, HR MS *m/z* found 355.1039 $[M + H]^+$ and $C_{16}H_{19}O_8{}^{18}O$ requires 357.1066, found 357.1082 $[M + H]^+$. The percentage of labelled scopolin was 12 %. The HPLC peak at 24.7 min, scopoletin, naturally occurring scopoletin $C_{10}H_9O_4$ requires 193.0495, HR MS *m/z* found 193.0494 $[M + H]^+$ and $C_{10}H_9O_3{}^{18}O$ requires 195.0538, found 195.0543 $[M + H]^+$. The percentage of labelled scopoletin was 4.8 %.

From the feeding experiments with $C^{18}O_2$-ferulic acid, the HPLC peak at 24.7 min, scopoletin, naturally occurring scopoletin $C_{10}H_9O_4$ requires 193.0495, HR MS *m/z* found 193.0489 $[M + H]^+$ and $C_{10}H_9O_3{}^{18}O$ requires 195.0538, found 195.0541 $[M + H]^+$. The percentage of labelled scopoletin was 3.7 %.

Feeding experiments in an atmosphere of $^{18}O_2$

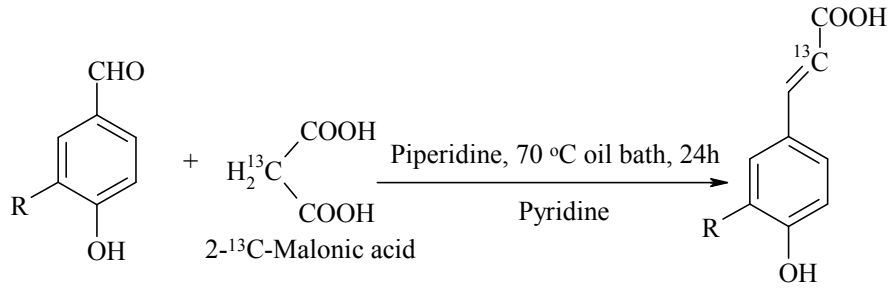
Cassava roots (0.38 kg, cv MNGA 2) were peeled (0.30 kg) and divided into three equal groups (85 g), the first and second groups were both fed with *E*-cinnamic-d_7 acid (20 mg) dissolved in aq. 4% Na_2CO_3 (2 ml); no exogenous substrate was added to the third group. The first and third groups were placed in a vacuum desiccator, which was immediately evacuated using a vacuubrand PC 2001 VARIO vacuum pump, with a vacuubrand cvc 2 pressure monitor, until a final pressure of 10 to 20 mbar was reached, and then was back-flushed with anhydrous nitrogen. This procedure was repeated twice more. After further evacuation, the desiccator was filled with anhydrous nitrogen only to a final pressure of 800 mbar and then taken to 1000 mbar with $^{18}O_2$ to afford an atmosphere $N_2/^{18}O_2$ approximately 4:1 v/v. The second group was allowed to deteriorate under the usual conditions in normal air. All three groups were allowed to deteriorate at 20 ^oC for 5 d, and then all were extracted following the general procedure. A representative sample of the EtOH extract (2 g) was then purified by HPLC. HR MS data of the HPLC peak, at 24.7 min, of $^{18}O_3$-scopoletin (fed in $^{18}O_2$, group 3) are listed in Table 4.13, and of $^{18}O_3$-scopoletin-d_3 (fed with cinnamic-d_7 acid in $^{18}O_2$, group 1) are shown in Fig. 4.26 and listed in Tables 4.13 and 4.14. The results from group 2 (fed with cinnamic-d_7 acid in normal air) were as reported above, and typical of Table 4.10.

4.3. Results and discussion

Synthesis of Isotopically Labelled Intermediates

Synthesis of p-coumaric-2-^{13}C acid, caffeic-2-^{13}C acid, and ferulic-2-^{13}C acid

p-Coumaric-2-^{13}C acid, caffeic-2-^{13}C acid and ferulic-2-^{13}C acid were synthesised via a Knoevenagel reaction (Robbins and Schmidt, 2004; Ji et al., 2005) by the nucleophilic addition of 2-^{13}C-malonic acid to the corresponding aldehyde (Fig. 4.7), and the three desired *E*-cinnamic acid derivatives were recrystallized from water. The crystal structure data (ORTEP diagram, for supplementary data see Appendix 1) of a ferulic-2-^{13}C acid single crystal X-ray is shown (Fig. 4.8).



R=H 4-Hydroxybenzaldehyde
R=OH 3,4-Dihydroxybenzaldehyde
R=OMe Vanillin

R=H *p*-Coumaric-2-^{13}C acid
R=OH Caffeic-2-^{13}C acid
R=OMe Ferulic-2-^{13}C acid

Fig. 4.7. Synthesis of labelled substituted cinnamic-2-^{13}C acids.

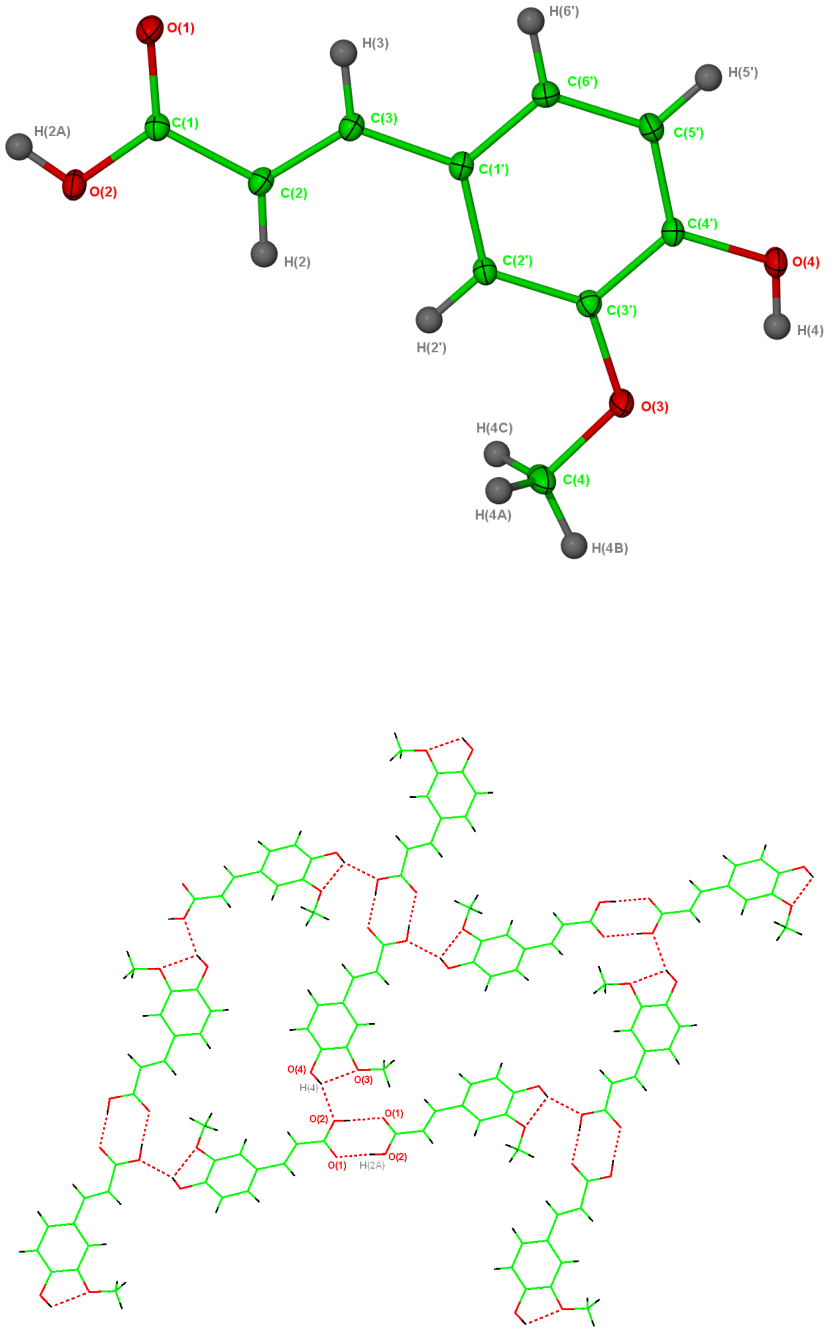


Fig. 4.8. ORTEP diagram X-ray crystallographic data from ferulic-2-^{13}C acid.

The chemical structures and therefore the position of the ^{13}C-labels were confirmed by ^{1}H and ^{13}C NMR spectroscopic data (Tables 4.1 and 4.2). The ^{1}H NMR spectra (Table 4.1, Fig. 4.9-4.12) showed H-2 doublet of doublets due to H-2-^{13}C-2 coupling ($^{1}J_{CH}$ = ~161) and *trans*-H-2-H-3 coupling ($^{3}J_{HH}$ = 16), H-3 is also a doublet of doublets including $^{2}J_{CH}$ = 3, and such *J* values (Hz) are consistent with the reported literature values (Williams and Fleming, 2008a). In all the ^{13}C NMR spectra, the signal for C-2 at δ 116 was approx. 90-fold more intense (99% enriched) than the natural abundance peaks for the other positions. This intense signal is superimposed on the doublet ($^{1}J_{CC}$ = 71) resulting from natural abundance ^{13}C at C-3 (or at C-1), and peaks at δ 171 and 147, corresponding to C-1 and C-3 respectively, are doublets ($^{1}J_{CC}$ = 74 or 71), due to this ^{1}J coupling. Peaks corresponding to C-2' and C-6' are doublets ($^{3}J_{CC}$ = 5) due to long-range coupling to the enriched C-2. Interestingly, quaternary C-1' was an apparent singlet ($^{2}J_{CC}$ < 1.5), due to a combination of being a quaternary carbon and being β-to the enriched ^{13}C-atom. That the $^{3}J_{CC}$ = 5 is greater than $^{2}J_{CC}$ < 1.5 is complex, and has literature precedent, e.g. in geraniol where $^{3}J_{CC}$ = 5 and $^{2}J_{CC}$ = 1.3 (Kozminski and Nanz, 1996), but this is not always the case as $^{2}J_{CC}$ can be equal to or slightly greater than $^{3}J_{CC}$, but always in the range 2-7 Hz (Reif et al., 1996) and such long-range ^{13}C-^{13}C coupling constants can now be detected at natural abundance with the appropriate pulse-sequence (Pham et al., 2005).

The ^{13}C NMR spectra (Table 4.2, Fig. 4.13-4.16) clearly show that introducing an *o*-phenolic group, from *p*-coumarate to caffeate, causes a shift of -12-16 ppm whilst a gain of 30 ppm is seen at the carbon which now carries the oxygen (relative to proton), and a shift of -8 ppm at the *para*-carbon. Therefore, we assign 149.5 (from 161.2) to 4' and 146.8 (from 116.8) to 3'. However, on converting the hydroxy into a methoxy, i.e. from caffeate to ferulate, the gain would be 30.2 ppm at the attached carbon, with a shift of -14.7 ppm at the *ortho*-carbons, according to tables (Williams and Fleming, 2008a). We cannot unambiguously assign 150.5 and 149.3 ppm to 3' and 4' respectively, but this order is consistent with tables, i.e. shifting further from TMS with a methoxy (relative to a hydroxy) and shifting slightly nearer (2 ppm) to TMS with an *o*-methoxy than with an *o*-hydroxy (Williams and Fleming, 2008a). These ^{1}H and ^{13}C NMR assignments are

also consistent with those few literature data which have been reported for cinnamic-2-^{13}C acid (Ma and Hayes, 2004) and ferulic-2-^{13}C acid (Bernards et al., 1999).

Table 4.1

^{1}H NMR (CD_3OD, 400 MHz) spectral data.

compound	***p*-coumaric-2-^{13}C acid**	**caffeic-2-^{13}C acid**	**ferulic-2-^{13}C acid**
position	**^{1}H multiplicity (*J*)**	**^{1}H multiplicity (*J*)**	**^{1}H multiplicity (*J*)**
2	6.28 dd (161, 16)	6.22 dd (160, 16)	6.31 dd (161, 16)
3	7.60 dd (16, 3)	7.53 dd (16, 3)	7.59 dd (16, 3)
2'	7.45 d (9)	7.04 d (2)	7.17 d (2)
3'	6.81 d (9)	-	-
5'	6.81 d (9)	6.78 d (8)	6.81 d (8)
6'	7.45 d (9)	6.94 dd (8, 2)	7.06 dd (8, 2)
OCH_3	-	-	3.89 s

Table 4.2

^{13}C NMR (CD_3OD, 100 MHz) spectral data (^{1}H broadband decoupled)

compound	***p*-coumaric-2-^{13}C acid**	**caffeic-2-^{13}C acid**	**ferulic-2-^{13}C acid**
position	**^{13}C multiplicity (*J*)**	**^{13}C multiplicity (*J*)**	**^{13}C multiplicity (*J*)**
1	171.0 d (74)	171.0 d (74)	171.0 d (74)
2	115.6 (99% enriched)	115.5 (99% enriched)	115.9 (99% enriched)
3	146.7 d (71)	147.0 d (71)	146.9 d (71)
1'	127.2	127.8	127.7
2'	131.1 d (5)	115.0 d (5)	111.6 d (5)
3'	116.8	146.8	150.5 or 149.3
4'	161.2	149.5	149.3 or 150.5
5'	116.8	116.5	116.2
6'	131.1 d (5)	122.8 d (5)	124.0 d (5)
OCH_3	-	-	56.4

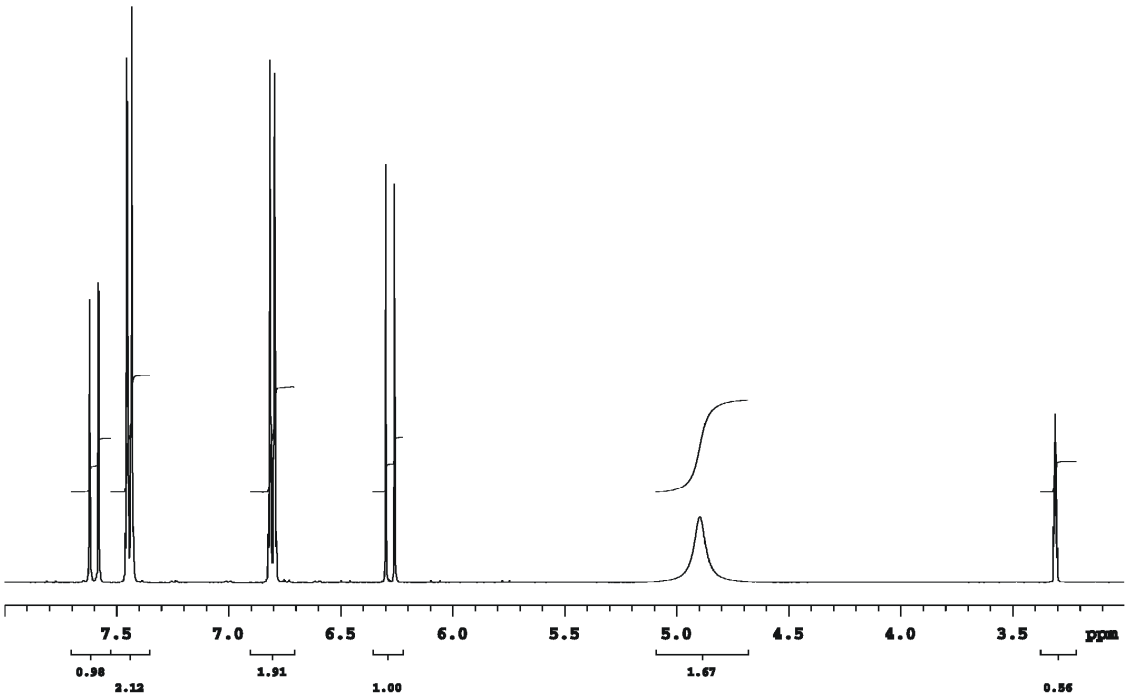


Fig. 4.9. ^{1}H NMR spectrum (CD_3OD, 400 MHz) of *p*-coumaric acid.

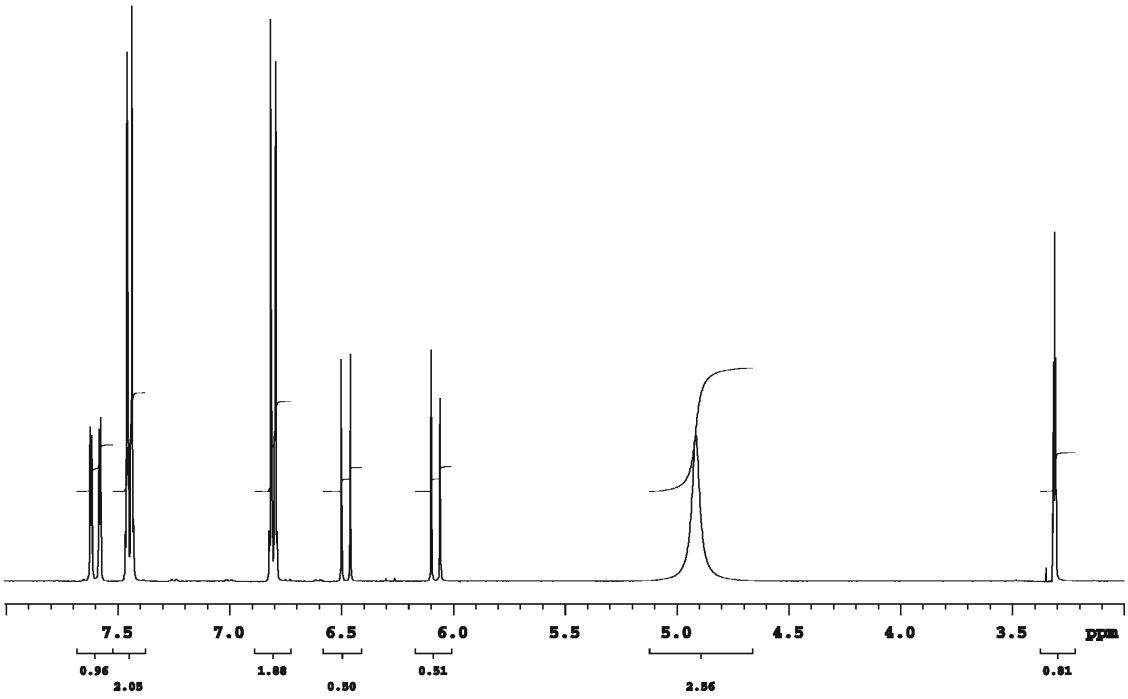


Fig. 4.10. ^{1}H NMR spectrum (CD_3OD, 400 MHz) of *p*-coumaric-2-^{13}C acid.

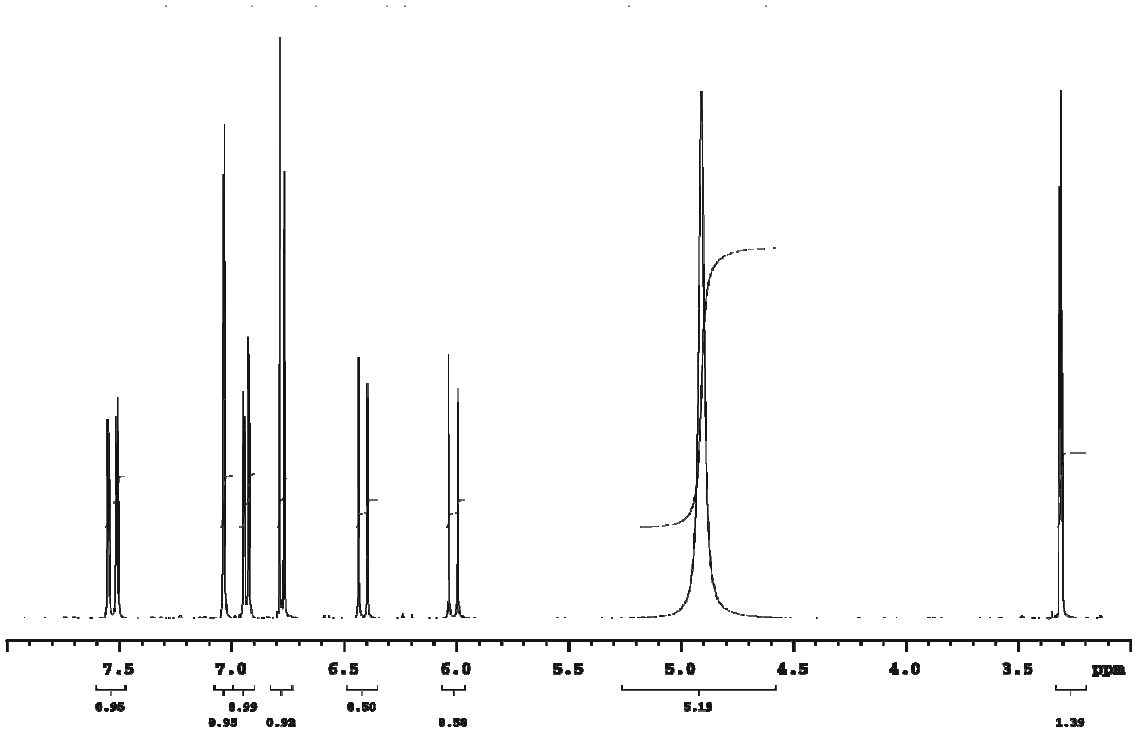


Fig. 4.11. ^{1}H NMR spectrum (CD_3OD, 400 MHz) of caffeic-2-^{13}C acid.

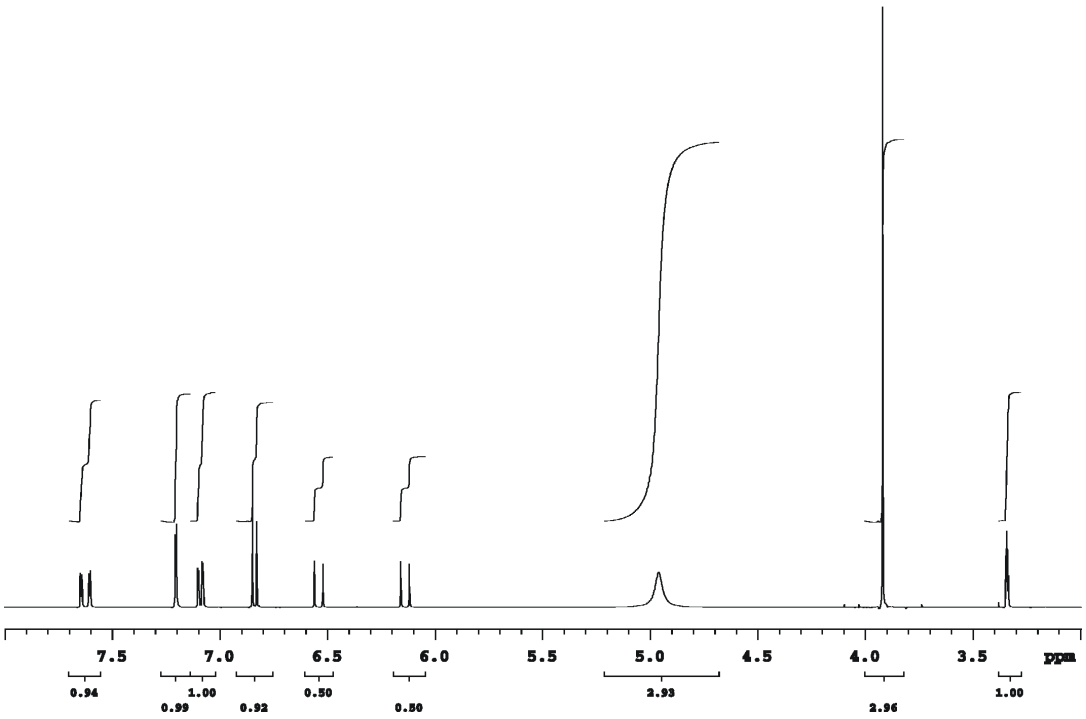


Fig. 4.12. ^{1}H NMR spectrum (CD_3OD, 400 MHz) of ferulic-2-^{13}C acid.

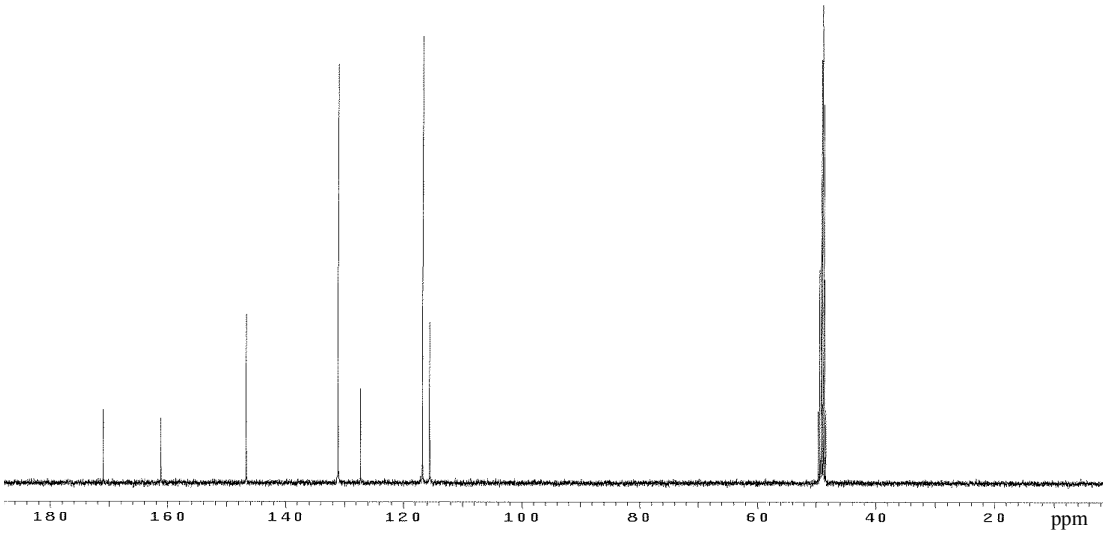


Fig. 4.13. ^{13}C NMR spectrum (CD_3OD, 100 MHz) of *p*-coumaric acid.

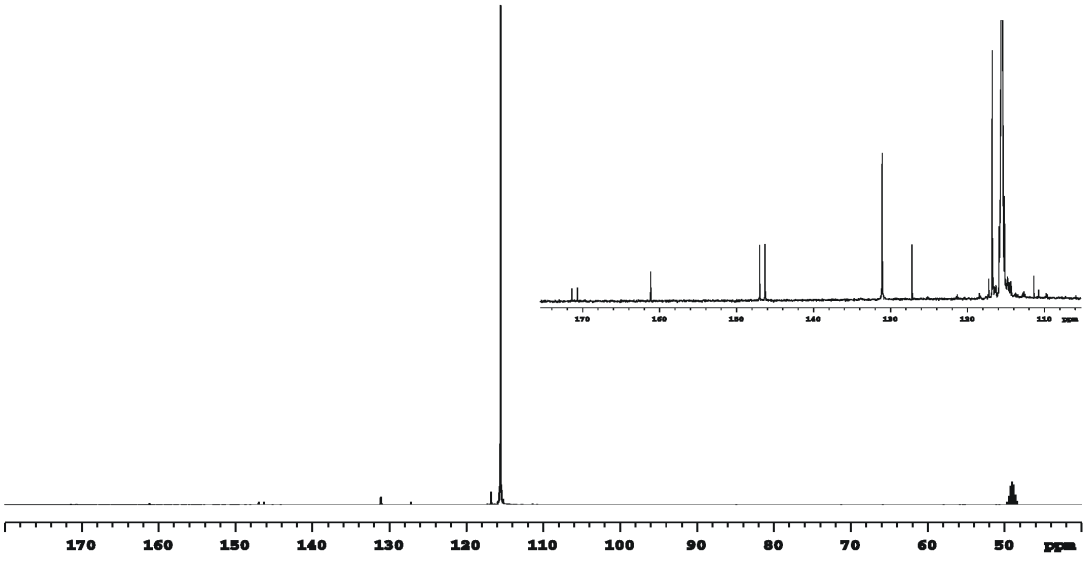


Fig. 4.14. ^{13}C NMR spectrum (CD_3OD, 100 MHz) of *p*-coumaric-2-^{13}C acid.

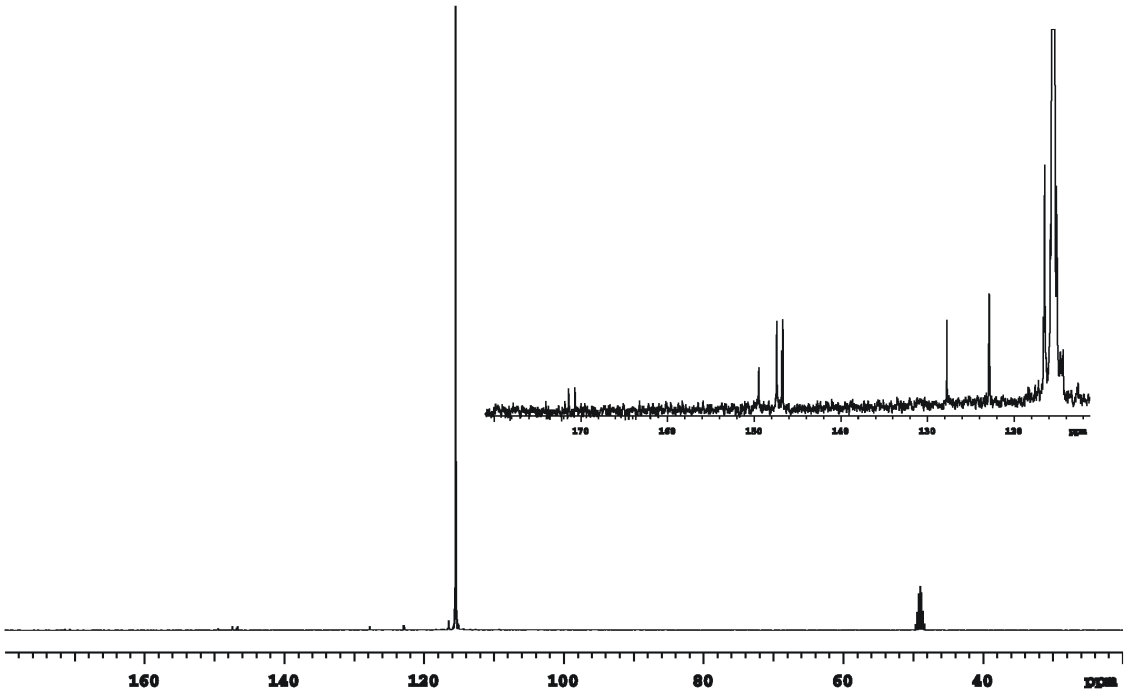


Fig. 4.15. ^{13}C NMR spectrum (CD_3OD, 100 MHz) of caffeic-2-^{13}C acid.

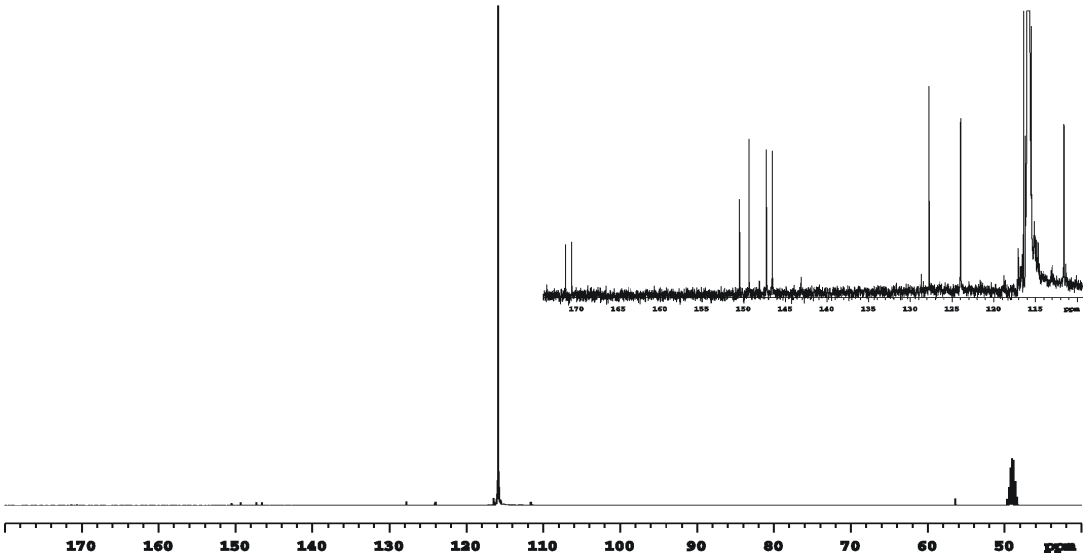


Fig. 4.16. ^{13}C NMR spectrum (CD_3OD, 100 MHz) of ferulic-2-^{13}C acid.

$^{18}O_1$-Enrichment of umbelliferone and esculetin

^{18}O-Enrichment of umbelliferone and esculetin with $H_2{}^{18}O$ is a practical method to introduce selectively only one label into these compounds (Fig. 4.17). This mechanism yields coumarins labelled only in the lactone carbonyl, the lactone ether and phenolic oxygen atoms are not exchanged. The percentage of ^{18}O-enrichment was monitored for umbelliferone after 3 h (8%), 4 d (61%) and 7 d (82%). Esculetin was ^{18}O-enriched for 7 d (82%), as this gave the highest percentage enrichment for umbelliferone.

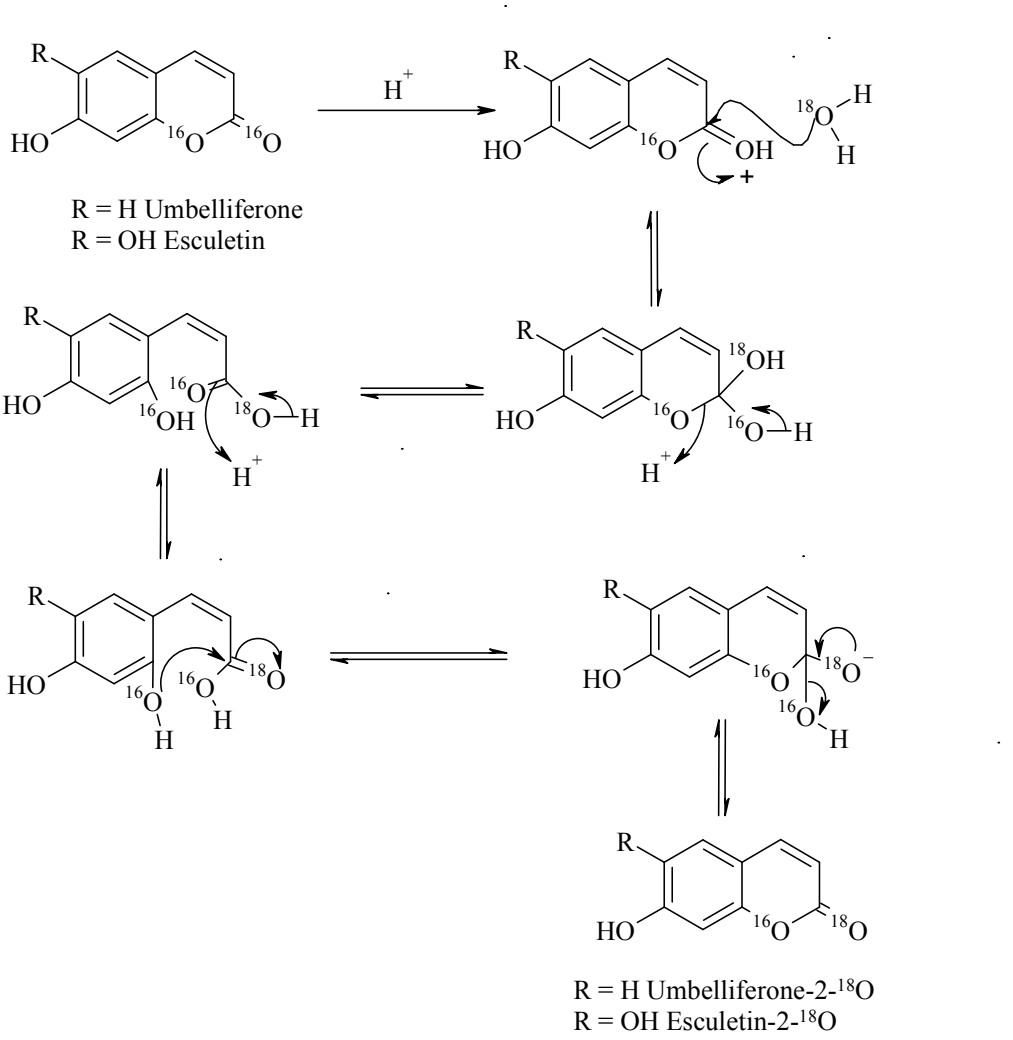
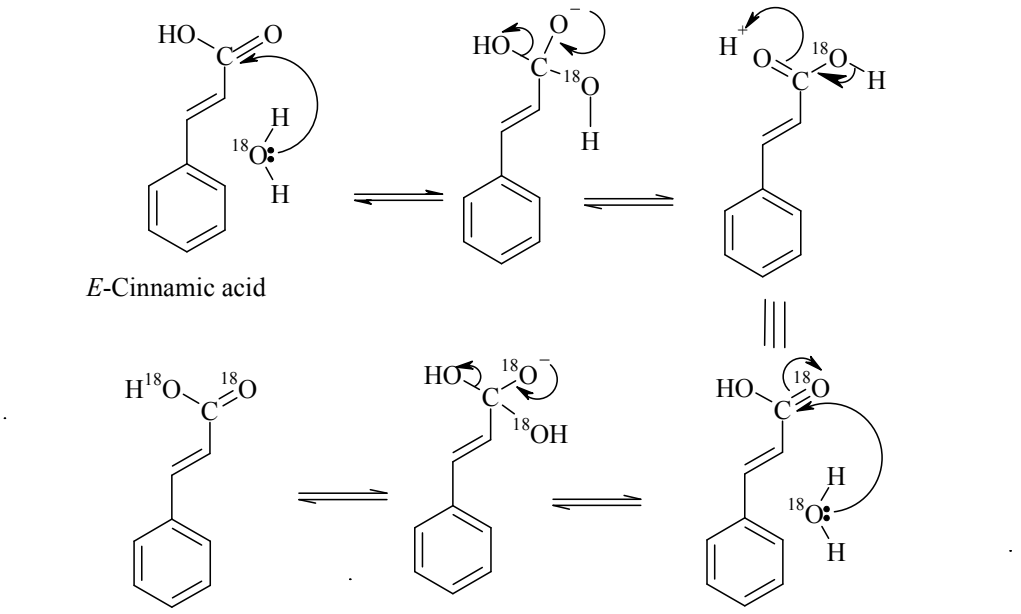


Fig. 4.17. Isotopic enrichment of umbelliferone or esculetin with ^{18}O by exchange with $H_2{}^{18}O$.

$C^{18}O_2$-Enrichment of E-cinnamic acid and ferulic acid

$C^{18}O_2$-Enriched *E*-cinnamic acid was synthesised by double exchange with $H_2{}^{18}O$ (Fig. 4.18) catalysed by conc. HCl. The percentage ^{18}O-enrichment was monitored over 3 h, 1, 2 and 4 d (Fig. 4.19) in order to optimise the yield of the $^{18}O_2$-enriched species (Table 4.3). $C^{18}O_2$-Enriched ferulic acid was then synthesised by the same method as *E*-cinnamic acid for 4 d.



$^{18}O_2$-Enriched *E*-cinnamic acid

Fig. 4.18. Scheme showing the mechanism of enrichment of *E*-cinnamic acid with ^{18}O by exchange with $H_2{}^{18}O$.

Table 4.3

HR ESI MS data of cinnamate labelling showing the percentage of ^{18}O-enriched *E*-cinnamic acid.

Duration of the reaction	**Percentage of unlabelled cinnamic acid % of *m/z* 149**	**Percentage of ^{18}O enrichment % of *m/z* 151**	**Percentage of $^{18}O_2$ enrichment % of *m/z* 153**
3 h	86.2	13.8	0
1 d	9.3	43.5	47.2
2 d	1.9	23.7	74.4
4 d	0	14.0	86.0

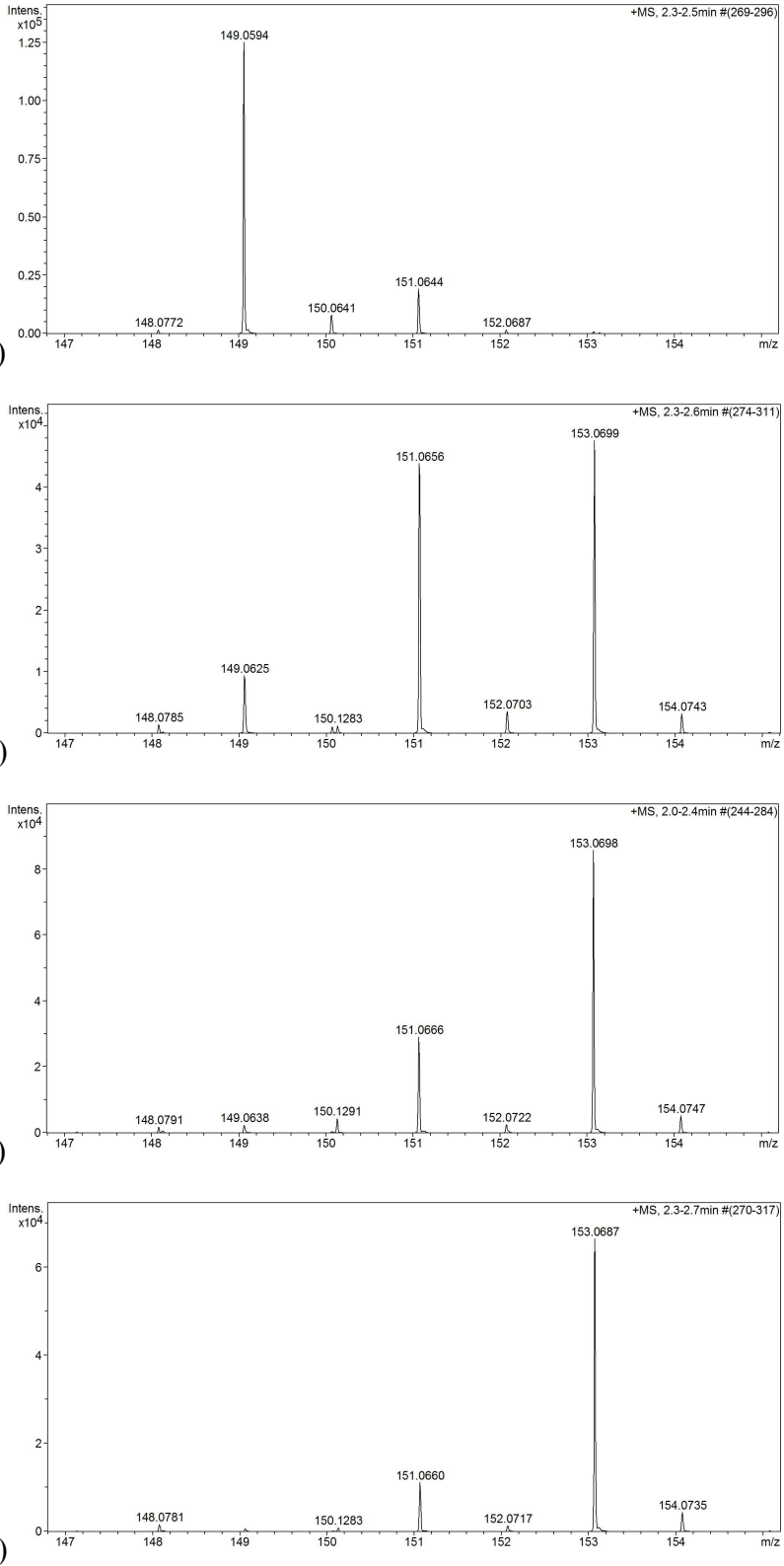


Fig. 4.19. a-d) HR Ms of cinnamic acid ^{18}O-enrichment was monitored over 3 h, 1, 2 and 4 d.

Biosynthetic Feeding Experiments

Feeding with p-coumaric-2-^{13}C acid, caffeic-2-^{13}C acid and ferulic-2-^{13}C acid

The relative importance of the pathways labelled 1, 2 and 3 (Fig. 4.4) was explored by direct feeding of cassava roots with the above three ^{13}C-labelled acids, which are later intermediates than cinnamic acid in the biosynthesis of scopoletin. *p*-Coumaric acid is an intermediate in all three proposed pathways, caffeic acid in pathways 2 and 3, while ferulic acid is in pathway 3 only. The percentage of labelled scopolin and scopoletin was calculated and is summarized in Tables 4.4 and 4.5. These results showed an increase in the percentage of labelled scopoletin (Fig. 4.20) and scopolin above the natural abundance for all three precursors. Ferulic-2-^{13}C acid incorporation confirms that pathway 3 is involved in the biosynthesis of scopoletin in cassava roots during PPD. However, as all three labelled precursors gave very similar enrichment of the product hydroxycoumarins, biosynthetic pathways 1 and 2 cannot be completely excluded. Experimental error was evaluated by calculating the ^{13}C-natural abundance of ten (unlabelled) scopoletin samples. The percentage of scopoletin containing 1 x ^{13}C was 7.7 ± 0.6% (s.d.) and that of 2 x ^{13}C-labelled scopoletin was 1.27 ± 0.05% (s.d.). This showed a low level of error in these experimental results.

Table 4.4

HR ESI MS data of the HPLC peak at 7.1 min showing the percentage of labelled scopolin.

Cassava cv MCOL 22 fed with	**Scopolin**	**Scopolin-3-^{13}C**	**% of Scopolin-3-^{13}C**	**% ^{13}C scopolin adjusted for nat. abundance**
Untreated (control)	355.1033	356.1073	13.9	-
p-Coumaric-2-^{13}C acid	355.1050	356.1087	31.2	17.3
Caffeic-2-^{13}C acid	355.1033	356.1079	28.2	14.3
Ferulic-2-^{13}C acid	355.1039	356.1068	27.5	13.6

Scopolin signifies *m/z* found of scopolin, $C_{16}H_{19}O_9$ $[M + H]^+$ (calcd. 355.1024)

Scopolin-3-^{13}C signifies *m/z* found of scopolin-3-^{13}C, $C_{15}{}^{13}CH_{19}O_9$ $[^{13}C, M + H]^+$ (calcd. 356.1057)

Table 4.5

HR ESI MS data of the HPLC peak at 24.7 min and the percentage of labelled scopoletin.

Cassava cv MCOL 22 fed with	**Scopoletin**	**Scopoletin-3-^{13}C**	**% of Scopoletin-3-^{13}C**	**% ^{13}C scopoletin adjusted for nat. abundance**
Untreated (control)	193.0487	194.0537	7.5	-
p-Coumaric-2-^{13}C acid	193.0488	194.0533	22.6	15.1
Caffeic-2-^{13}C acid	193.0489	194.0529	19.1	11.6
Ferulic-2-^{13}C acid	193.0487	194.0529	23.0	15.5

Scopoletin signifies *m/z* found of scopoletin, $C_{10}H_9O_4$ $[M + H]^+$ (calcd. 193.0495)

Scopoletin-3-^{13}C signifies *m/z* found of scopoletin-3-^{13}C, $C_9{}^{13}CH_9O_4$ $[^{13}C, M + H]^+$ (calcd. 194.0529)

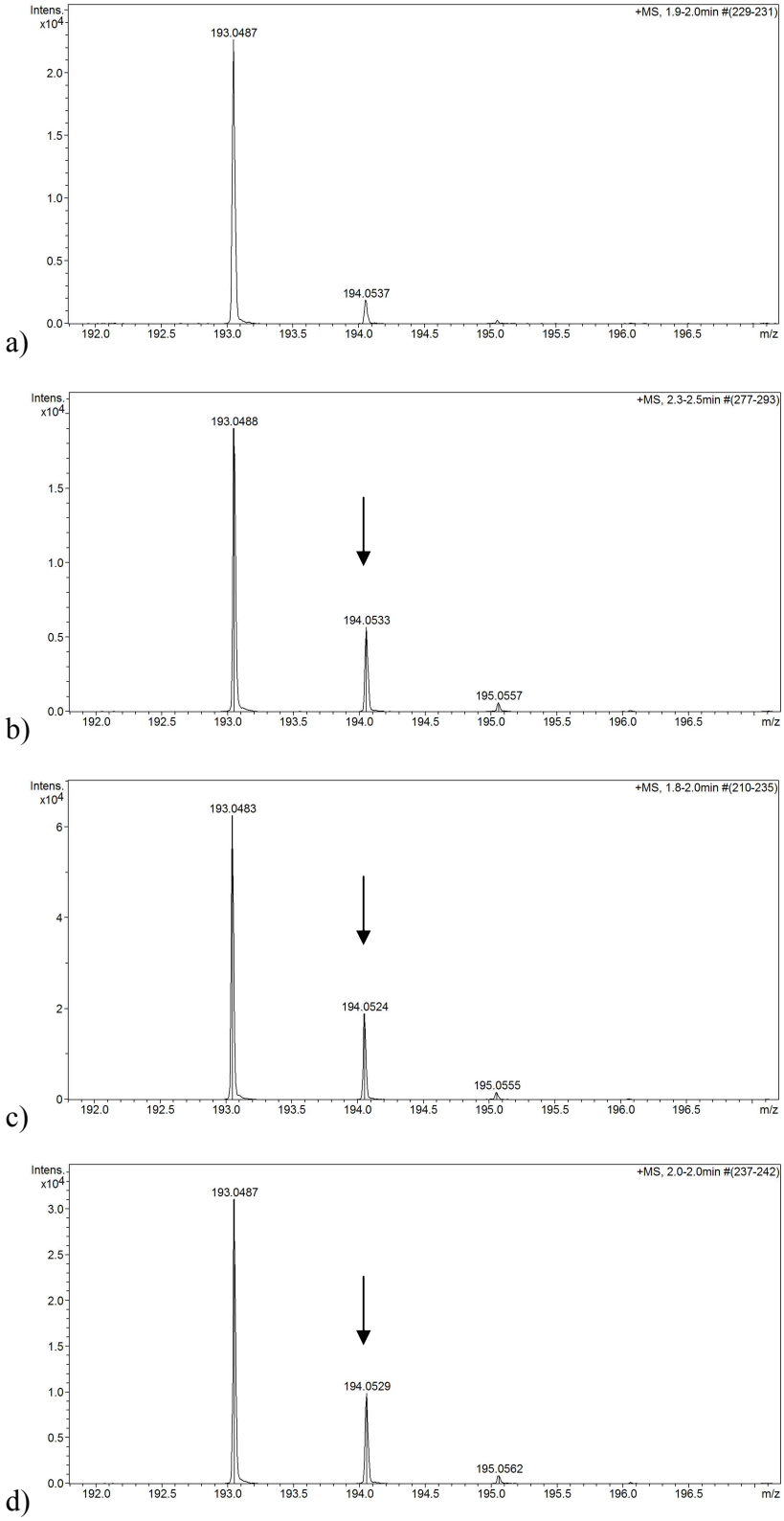


Fig. 4.20. a) HR MS of scopoletin from deteriorated root (control); b-d) HR MS of scopoletin after feeding with *p*-coumaric-2-^{13}C, caffeic -2-^{13}C and ferulic-2-^{13}C respectively.

In another feeding experiment with $C^{18}O_2$-enriched *E*-cinnamic acid, *p*-coumaric-2-^{13}C acid, caffeic-2-^{13}C acid and ferulic-2-^{13}C acid the roots did not go black over the incubation period; the PPD process appeared to have failed and the amount of scopoletin recovered by HPLC was small. The percentage label scopoletin from each of the precursors $C^{18}O_2$-enriched *E*-cinnamic acid, *p*-coumaric-2-^{13}C acid, caffeic-2-^{13}C acid and ferulic-2-^{13}C acid was 3.9 % 29.9 %, 38.9 %, 37.7 % respectively. From these results it is concluded that in this experiment cinnamate-4'-hydroxylase showed only low activity, that later enzymes in the pathway were active, and that the high incorporation of label from later precursors (*p*-coumaric-2-^{13}C acid, caffeic-2-^{13}C acid and ferulic-2-^{13}C acid) was a result of a low flux of natural abundance intermediated through the pathway. This experiment confirmed the correlation between the black colour of the roots as a manifestation of PPD and the accumulation scopoletin in the roots during PPD. The cause of the low activity of cinnamate-4'-hydroxylase is unknown but down regulation or inhibition of this gene could well produce a similar arrest of PPD.

Feeding with ^{18}O-umbelliferone and ^{18}O-esculetin compared with ferulic-2-^{13}C acid

The relative importance of these three possible pathways (Fig. 4.4) was then further explored by feeding cassava roots with the above three compounds, later intermediates in these three pathways where umbelliferone is in pathway 1, esculetin is in both pathways 1 and 2, but ferulic acid is only in pathway 3. The percentage of labelled scopolin, scopoletin and esculin is summarized in Tables 4.6, 4.7 and 4.8. These results showed an increase in the percentage of labelled scopoletin (Fig. 4.21) and scopolin above the natural abundance for ferulic-2-^{13}C acid, together with a relatively small increase in the percentage of labelled scopoletin and scopolin above the natural abundance for umbelliferone and esculetin, which we interpret as confirmation that pathway 3 (Fig. 4.22) is the main pathway for the biosynthesis of scopoletin and scopolin in cassava roots during PPD.

Table 4.6

HR MS data of the HPLC peak at 7.1 min showing the percentage of labelled scopolin.

Feeding cassava cv MNGA 19 with	**Scopolin**	**Isotope peak monitored**	***m/z* found of labelled scopolin (calcd.)**	**% of total labelled scopolin**	**% of enriched scopolin**
Ferulic-2-^{13}C acid	355.1012	$C_{15}{}^{13}CH_{19}O_9$	356.1053 (356.1057)	33.3	19.4
Umbelliferone-2-^{18}O	355.1036	$C_{16}H_{19}O_8{}^{18}O$	357.1396 (357.1066)	9.4	4.8
Esculetin-2-^{18}O	355.1024	$C_{16}H_{19}O_8{}^{18}O$	357.1102 (357.1066)	8.3	3.7

Scopolin signifies *m/z* found of scopolin, $C_{16}H_{19}O_9$ $[M + H]^+$ (calcd. 355.1024)

Table 4.7

HR MS data of the HPLC peak at 24.7 min showing the percentage of labelled scopoletin.

Feeding cassava cv MNGA 19 with	**Scopoletin**	**Isotope peak monitored**	***m/z* found of labelled scopoletin (calcd.)**	**% of total labelled scopoletin**	**% of enriched scopoletin**
Ferulic-2-^{13}C acid	193.0499	$C_9{}^{13}CH_9O_4$	194.0532 (194.0529)	15.7	8.2
Umbelliferone-2-^{18}O	193.0491	$C_{10}H_9O_3{}^{18}O$	195.0524 (195.0538)	1.7	0.4
Esculetin-2-^{18}O	193.0493	$C_{10}H_9O_3{}^{18}O$	195.0493 (195.0538)	1.6	0.3

Scopoletin signifies *m/z* found of scopoletin, $C_{10}H_9O_4$ $[M + H]^+$ (calcd. 193.0495)

It is noteworthy that the accurate mass of the [M+2] isotope peak for scopolin (Table 4.6) obtained from the experiment with ^{18}O-labelled umbelliferone and esculetin did not agree with calculated value for $C_{16}H_{19}O_8{}^{18}O$. Thus, the peak observed at this

mass does not arise entirely from ^{18}O-labelled scopolin, which would reduce the apparent enrichment observed. A similar discrepancy was observed in the accurate mass of scopoletin derived from ^{18}O-labelled esculetin (Table 4.7).

Table 4.8
HR MS data of the HPLC peak at 6.0 min.

Feeding cassava cv MNGA 19 with	**Esculin**	***m/z* found of esculin-2-^{18}O**	**% of labelled esculin**
Umbelliferone-2-^{18}O	341.0864	343.0921	19.8
Esculetin-2-^{18}O	341.0871	343.0907	67.7

Esculin signifies *m/z* found of esculin, $C_{15}H_{17}O_9$ $[M + H]^+$ (calcd. 341.0867)
Esculin-2-^{18}O signifies *m/z* found of esculin-2-^{18}O, $C_{15}H_{17}O_8{}^{18}O$
$[^{18}O, M + H]^+$ (calcd. 343.0910)

The high percentage of label incorporation into the glucoside esculin (Table 4.8) when umbelliferone-2-^{18}O (gave 20% labelled esculin) and esculetin-2-^{18}O (gave 68% labelled esculin) were fed to cassava roots showed that glucosyltransferase activity is high during PPD. This merits further investigation and is part of the subject of Chapter 5 below.

From the incorporation of umbelliferone into esculin in good yield (Table 4.8), it is concluded that pathway 1 (Fig. 4.22) operates for the biosynthesis of esculetin with only a small level of *O*-methylation to scopoletin. While pathway 3 is the major pathway for the biosynthesis of scopoletin and scopolin. The key first intermediate in pathway 1, 2',4'-dihydroxycinnamic acid, was investigated in a competition experiment with *E*-cinnamic-d_7 acid in a further attempt to elucidate the biosynthetic pathway.

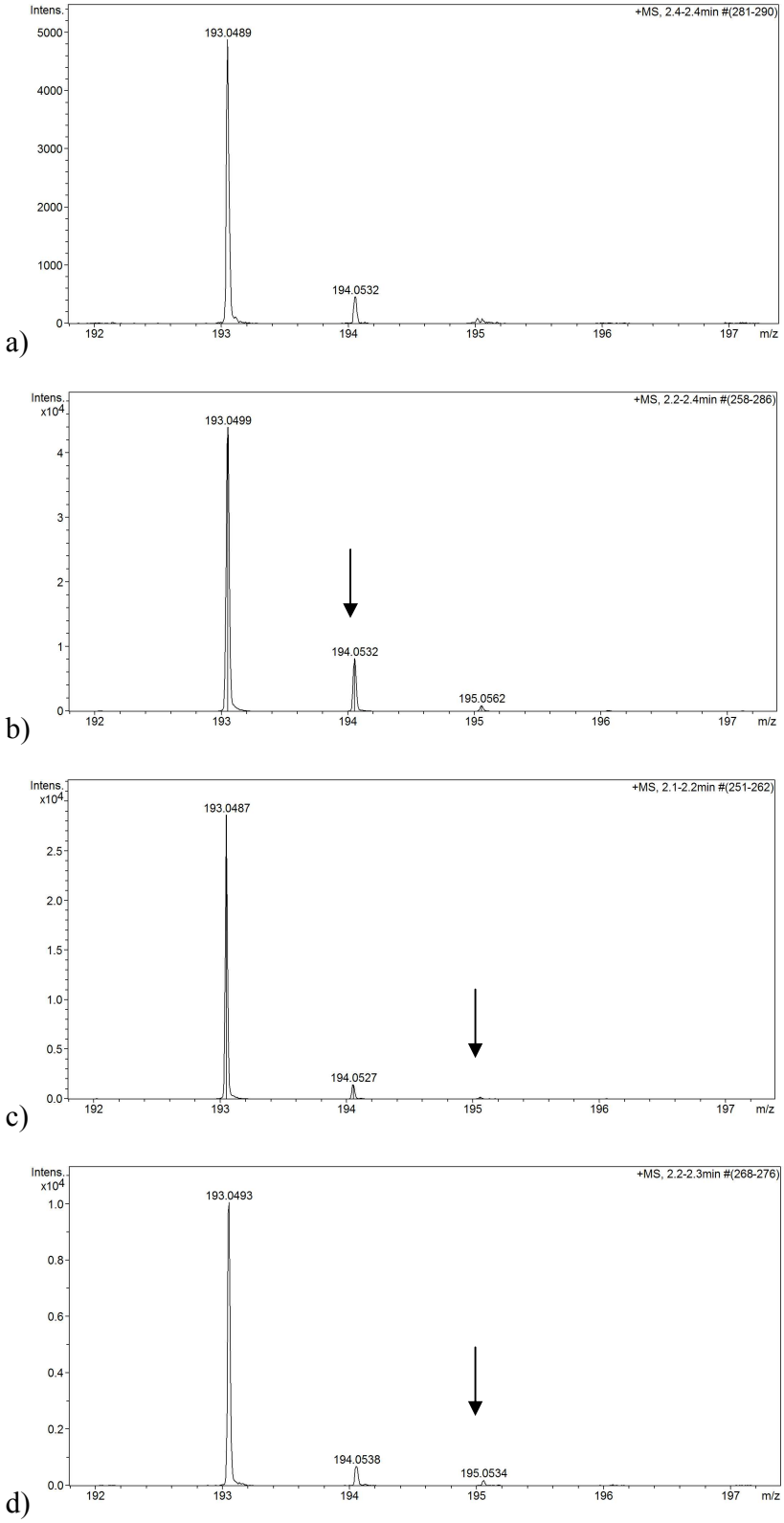


Fig. 4.21 a) HR MS of scopoletin from deteriorated root (control); b-d) HR MS of scopoletin after feeding with ferulic-2-^{13}C acid, umbelliferone-2-^{18}O and esculetin-2-^{18}O.

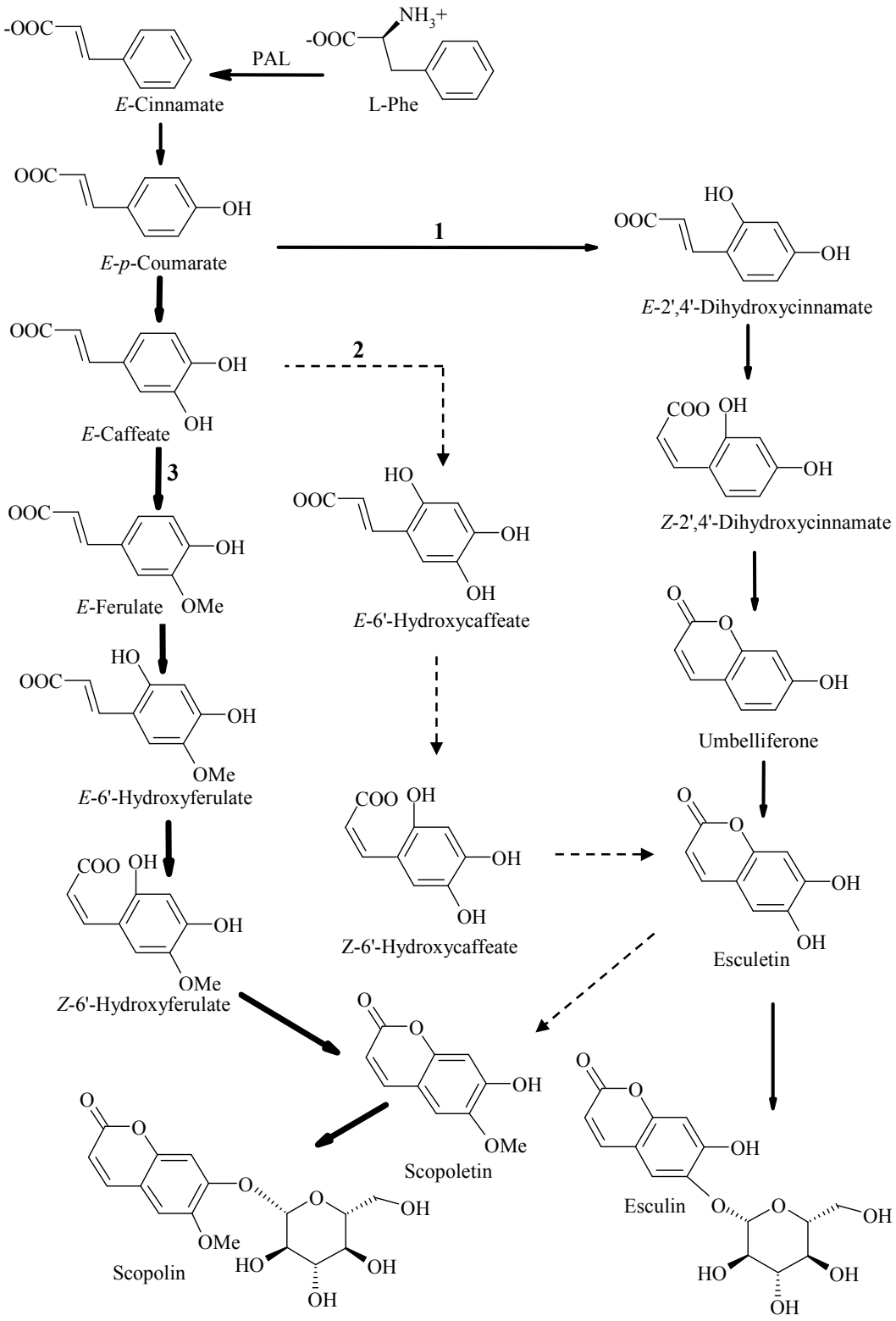


Fig. 4.22. Results of feeding experiments supporting pathway 3 as the major pathway for the biosynthesis of scopoletin in cassava roots during PPD.

Competition of deuterium labelled cinnamic acid and unlabelled intermediates feeding experiments

Competition experiments were carried out by feeding the root samples with *E*-cinnamic-d_7 acid alone as a control, and with *E*-cinnamic-d_7 together with unlabelled 2',4'-dihydroxycinnamic, ferulic or caffeic acids. It was postulated (Brown et al., 1984; Brown et al., 1988; Kai et al., 2006; Kai et al., 2008) that if any of these unlabelled compounds that are intermediates in the biosynthesis of scopoletin, they would compete with *E*-cinnamic-d_7 acid and decrease the amount of deuterium labelled scopoletin produced. The results of these experiments are summarized in Tables 4.9 and 4.10.

Table 4.9

HR MS data of the HPLC peak at 7.1 min showing the percentage of labelled scopolin.

Feeding cassava cv MCOL 22 with	**Scopolin**	**Scopolin-d_3**	**% of scopolin-d_3**
Cinnamic-d_7 acid	355.1027	358.1222	7.8
Cinnamic-d_7 + ferulic acids	355.1037	358.1224	8.1
Cinnamic-d_7 + caffeic acids	355.1038	358.1217	14.1
Cinnamic-d_7 + 2',4'-dihydroxycinnamic acids	355.1037	358.1224	7.5

Scopolin signifies *m/z* found of scopolin, $C_{16}H_{19}O_9$ $[M + H]^+$ (calcd. 355.1024)
Scopolin-d_3 signifies *m/z* found of scopolin-d_3, $C_{16}H_{16}D_3O_9$ $[D_3, M + H]^+$ (calcd. 358.1212)

Table 4.10

HR MS data of the HPLC peak at 27.4 min showing the percentage of labelled scopoletin.

Feeding cassava cv MCOL 22 with	**Scopoletin**	**Scopoletin-d_3**	**% of scopoletin-d_3**
Cinnamic-d_7 acid	193.0502	196.0694	5.8
Cinnamic-d_7 + ferulic acids	193.0489	196.0687	6.0
Cinnamic-d_7 + caffeic acids	193.0496	196.0692	8.3
Cinnamic-d_7 + 2',4'-dihydroxycinnamic acids	193.0500	196.0688	4.3

Scopoletin signifies *m/z* found of scopoletin, $C_{10}H_9O_4$ $[M + H]^+$ (calcd. 193.0495)

Scopoletin-d_3 signifies *m/z* found of scopoletin-d_3, $C_{10}H_6D_3O_4$ $[D_3, M + H]^+$ (calcd. 196.0684)

A further competition experiment was carried out between labelled *E*-cinnamic-2',3',4',5',6'-d_5 acid and umbelliferone or esculetin which are in the same pathway as 2',4'-dihydroxycinnamate (pathway 1, Fig. 4.22). The results of these experiments are summarized in Tables 4.11 and 4.12. Only minor reductions of enrichment of labelled scopolin and scopoletin were observed with 2',4'-dihydroxycinnamate and umbelliferone, confirming a minor role for this pathway in the biosynthesis of scopoletin. The presence of unlabelled caffeic acid or esculetin resulted in an unexpected increase in the percentage of labelling in scopoletin and scopolin. It is possible that this effect arises from inhibition of PAL (Sato et al., 1982) reducing the flux of unlabelled substrate through the pathway.

Table 4.11

HR ESI MS data of peak 1 at 7.1 min showing the percentage of labelled scopolin.

Feeding cassava cv MCOL 22 with	**Scopolin**	**Scopolin-d_2**	**% of scopolin-d_2**
Cinnamic-d_5 acid	355.1018	357.1150	16.8
Cinnamic-d_5 acid + umbelliferone	355.1036	357.1174	14.0
Cinnamic-d_5 acid + esculetin	355.1028	357.1150	25.8

Scopolin signifies *m/z* found of scopolin, $C_{16}H_{19}O_9$ $[M + H]^+$ (calcd. 355.1024)

Scopolin-d_2 signifies *m/z* found of scopolin-d_2, $C_{16}H_{17}D_2O_9$ $[D_2, M + H]^+$ (calcd. 357.1149)

Table 4.12

HR ESI MS data of peak 2 at 24.7 min showing the percentage of labelled scopoletin.

Feeding cassava cv MCOL 22 with	**Scopoletin**	**Scopoletin-d_2**	**% of scopoletin-d_2**
Cinnamic-d_5 acid	193.0501	195.0633	8.1
Cinnamic-d_5 acid + umbelliferone	193.0502	195.0623	4.8
Cinnamic-d_5 acid + esculetin	193.0503	195.0631	14.9

Scopoletin signifies *m/z* found of scopoletin, $C_{10}H_9O_4$ $[M + H]^+$ (calcd. 193.0495)

Scopoletin-d_2 signifies *m/z* found of scopoletin-d_2, $C_{10}H_7D_2O_4$ $[D_2, M + H]^+$ (calcd. 195.0621)

Feeding experiments with $C^{18}O_2$-labelled E-cinnamic and ferulic acids

All the hydroxycoumarins biosynthetic pathways ultimately require closure of the lactone ring to form the coumarin ring system. If the biosynthesis of scopoletin occurs via a spirodienone intermediate, then feeding experiments with $C^{18}O_2$-enriched *E*-cinnamic acid would yield $^{18}O_2$-scopoletin enriched in both the lactone ether and carbonyl oxygen atoms, whereas biosynthesis through 2'-hydroxylation of 4'-hydroxycinnamic acid would yield $^{18}O_1$-scopoletin enriched only in the carbonyl oxygen atom (Fig. 4.23). Both ^{18}O-double labelled *E*-cinnamic and ferulic acids were

incorporated only into scopoletin-2-^{18}O and scopolin-2-^{18}O (Fig. 4.24), i.e. scopoletin and scopolin only enriched in the carbonyl oxygen. No HR MS peaks corresponding to $C_{10}H_9{}^{16}O_2{}^{18}O_2$ or $C_{16}H_{19}{}^{16}O_7{}^{18}O_2$ were found; this is evidence that their biosynthesis in cassava roots during PPD occurs via *o*-hydroxylation and not via a spirodienone intermediate where both ^{18}O-atoms would have been incorporated in the final product. The accurate mass of the [M+2] isotope peak for scopoletin obtained from feeding experiments with $C^{18}O_2$-labelled *E*-cinnamic and ferulic acids is (found) 195.0541, and again 195.0543, agreeing with the value calculated for $C_{10}H_9{}^{16}O_3{}^{18}O$ (195.0538). The [M+2] peak (Fig. 4.24) is not due to the presence of two ^{13}C-isotopes at natural abundance which is calculated as 195.0562 for $C_8{}^{13}C_2H_9{}^{16}O_4$. From Einstein's Theory of Relativity, $E = mc^2$, some of the energy used for holding the extra neutrons in the nucleus has come from the mass. Thus, although at a superficial level adding two neutrons to make ^{18}O from ^{16}O looks like adding any other two neutrons e.g. incorporating 2 x ^{13}C or 2 x D, it is not the same mass gain, and this small, but measurable difference is called the mass defect.

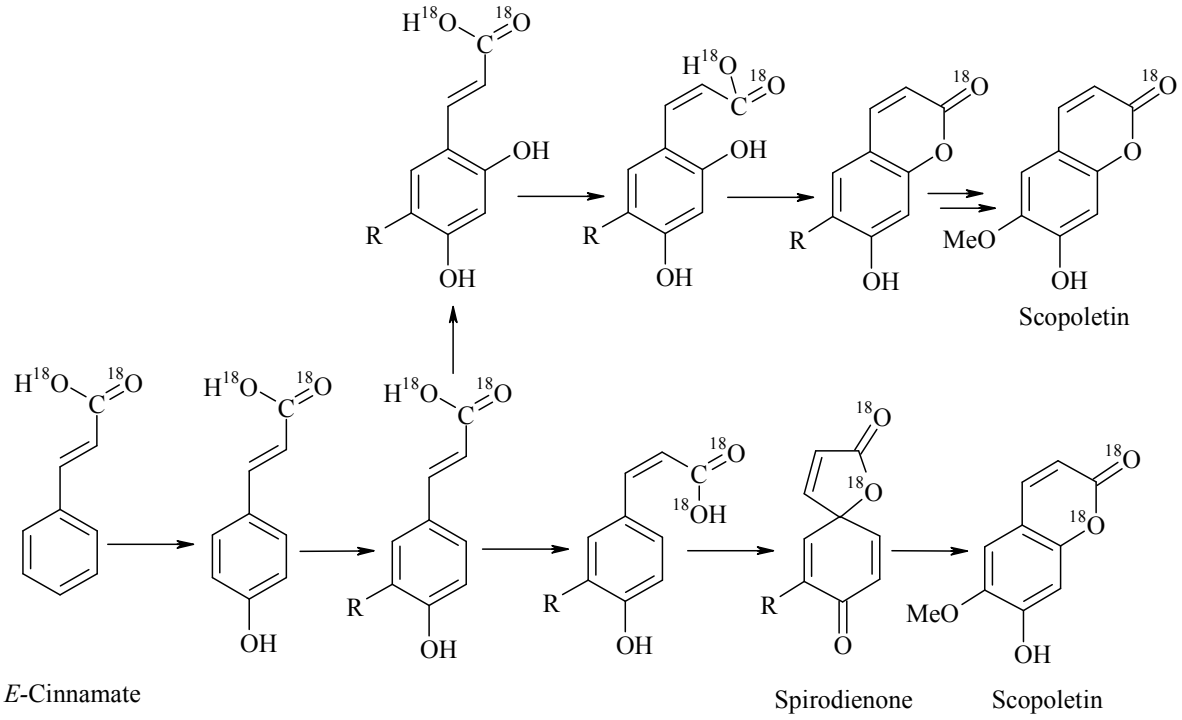
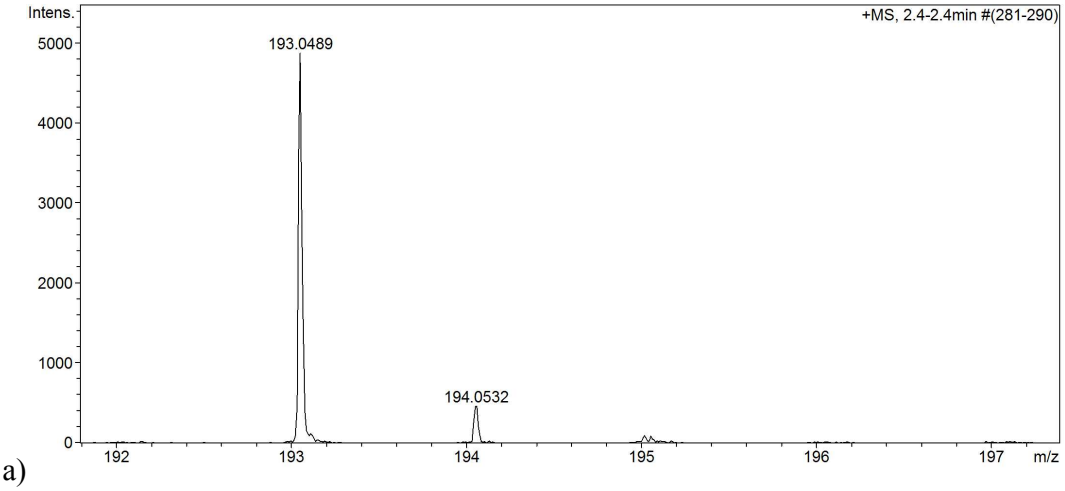
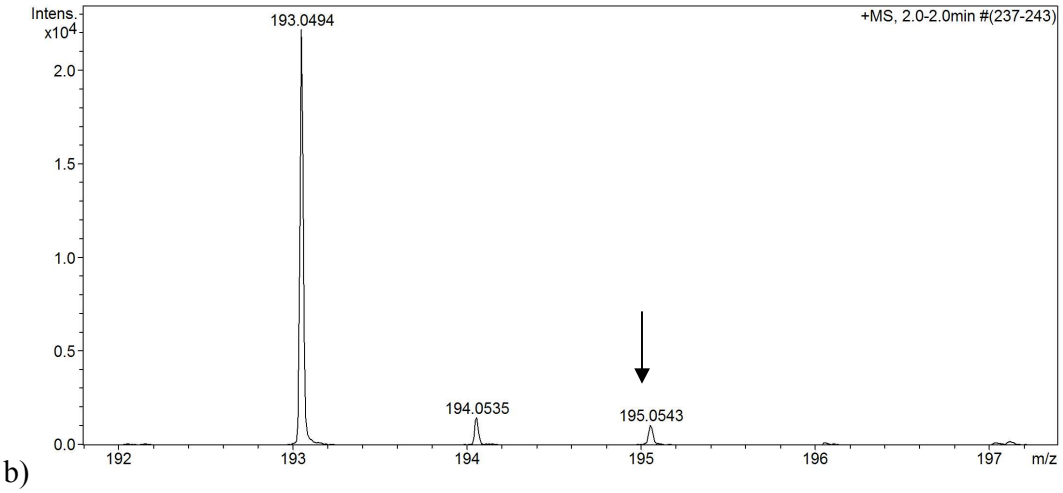


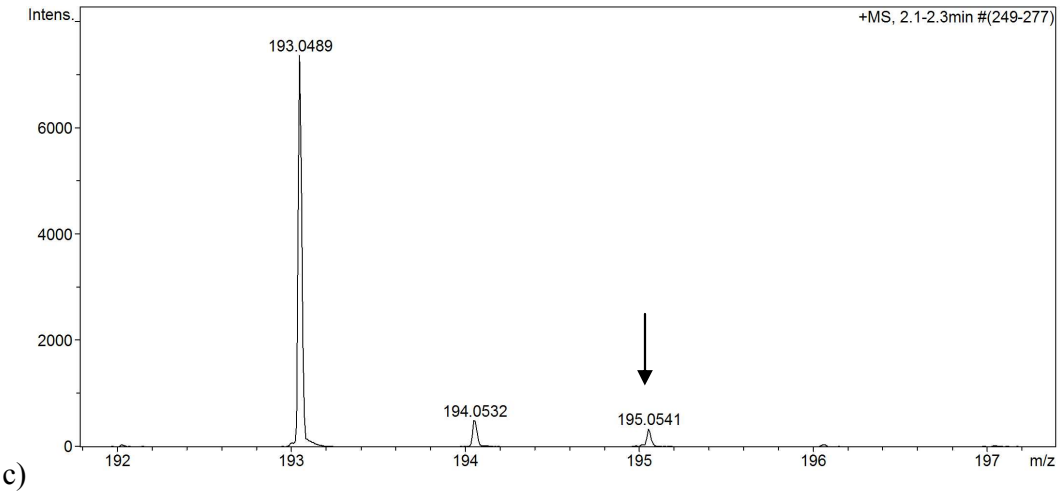
Fig. 4.23. Pathways from feeding $C^{18}O_2$-enriched *E*-cinnamic acid depending on the route of the biosynthesis of scopoletin in cassava roots.



a)



b)



c)

Fig. 4.24. a) HR MS of scopoletin from deteriorated root (control); b,c) HR MS of scopoletin after feeding with $C^{18}O_2$-labelled *E*-cinnamic and ferulic acids.

Feeding experiments in an atmosphere of 20% $^{18}O_2$

It is possible that, the ^{18}O-double labelled cinnamic acid could lose one ^{18}O-oxygen during the pathway to scopoletin by conversion into its CoA ester or into the shikimate or quinate ester;(Kai et al., 2006) such esters have been shown to be involved during the insertion of the 3-hydroxy group in several species. Although ferulate is a later component of the pathway, the involvement of such an ester of ferulate in the *E-Z*-isomerisation step cannot be ruled out. In order to confirm our interpretation of the above results, feeding experiments in an atmosphere rich in $^{18}O_2$ have been designed and carried out using a vacuum desiccator evacuated to 10 to 20 mbar (1 mbar is 100 Pa, 100 N/m^2) and then filled with anhydrous nitrogen only to a final pressure of 800 mbar and finally taken to 1000 mbar with $^{18}O_2$ to afford an atmosphere $N_2/^{18}O_2$ approximately 4:1 v/v.

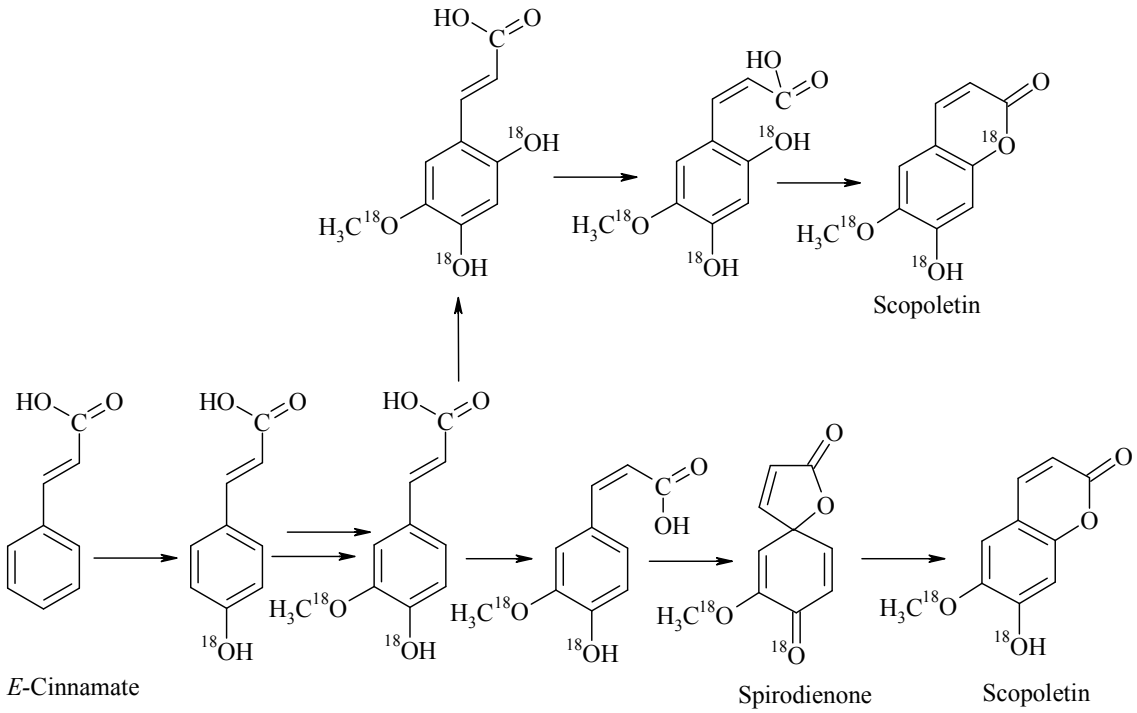


Fig. 4.25. Possible biosynthetic routes to scopoletin in cassava roots in $^{18}O_2$.

It has been established that both 4'-hydroxylation and 3'-hydroxylation of cinnamate involve cytochrome P450 enzymes (Bourgaud et al., 2006) and that the inserted oxygen atom is derived from molecular oxygen. 2'-Hydroxylation is thought to involve a Fe(II)- and 2-oxoglutarate-dependent dioxygenase (Kai et al., 2008) which also utilises molecular oxygen. Thus, if the biosynthesis of scopoletin occurs via a spirodienone intermediate, feeding experiments in $^{18}O_2$ would yield $^{18}O_2$-enriched scopoletin, doubly labelled in the methoxy and hydroxy oxygen groups, whereas biosynthesis through *o*-hydroxylation would yield $^{18}O_3$-enriched scopoletin, labelled in the methoxy, hydroxy and lactone ether oxygen atoms (Fig. 4.25). The results of scopoletin biosynthesis in $^{18}O_2$ and of feeding cassava roots with cinnamic-d_7 acid in $^{18}O_2$ are summarized in Tables 4.13 and 4.14 and two representative HR MS are shown in Fig. 4.26.

Table 4.13

HR ESI MS data of scopoletin showing the percentage of labelled scopoletin when cassava cv MNGA 2 was stored in 20% $^{18}O_2$.

Scopoletin and its isotope peak monitored $[M+H]^+$	***m/z* found of labelled scopoletin**	**Calcd**	**% of scopoletin isotopomers**
$C_{10}H_9{}^{16}O_4$	193.0496	193.0495	8.8
$C_{10}H_9O^{16}O_3{}^{18}O$	195.0544	195.0538	10.7
$C_{10}H_9{}^{16}O_2{}^{18}O_2$	197.0581	197.0580	33.3
$C_{10}H_9{}^{16}O^{18}O_3$	199.0624	199.0623	47.2

Table 4.14

HR ESI MS data of scopoletin showing the percentage of labelled scopoletin when cassava cv MNGA 2 was fed with cinnamic-d_7 acid in $^{18}O_2$.

Scopoletin and its isotope peak monitored $[M+H]^+$	***m/z* found of labelled scopoletin**	**Calcd**	**% of scopoletin isotopomers**
$C_{10}H_9{}^{16}O_4$	193.0502	193.0495	2.3
$C_{10}H_9{}^{16}O_3{}^{18}O$	195.0571	195.0538	1.8
$C_{10}H_9{}^{16}O_2{}^{18}O_2$	197.0606	197.0580	4.5
$C_{10}H_9{}^{16}O{}^{18}O_3$	199.0656	199.0623	25.2
$C_{10}H_6D_3{}^{16}O_4$	196.0684	196.0684	3.2
$C_{10}H_6\ D_3{}^{16}O_3{}^{18}O$	198.0719	198.0726	9.9
$C_{10}H_6\ D_3{}^{16}O_2{}^{18}O_2$	200.0768	200.0769	24.1
$C_{10}H_6\ D_3{}^{16}O{}^{18}O_3$	202.0804	202.0811	29.0

These results (Fig. 4.26) unequivocally show that the major isotopic peak was $^{18}O_3$-enriched scopoletin, and thus the major pathway in the biosynthesis of scopoletin in cassava roots during PPD is through *o*-hydroxylation not via a spirolactone-dienone intermediate. Although there is $^{18}O_2$-enriched scopoletin, this is mainly due to the presence of a low percentage of residual air (i.e. of $^{16}O_2$) trapped in the plant material.

This was confirmed by the presence of small amounts of both unlabelled scopoletin (2.3%) and $^{18}O_1$-enriched scopoletin (1.8%, Table 4.14). Feeding cassava roots with cinnamic-d_7 acid in $^{18}O_2$ resulted in the formation of a small amount of $^{16}O_4$-scopoletin-d_3 (3.2%) and there was a similar pattern of ^{18}O-labelling superimposed on scopoletin-d_3. Thus, as well as the definitive $[M+6]^+$ peak for $^{18}O_3$-scopoletin $[M + H]^+$ (Fig. 4.26a), the corresponding $[M+6]^+$ peak for $^{18}O_3$-scopoletin-d_3 $[M + Na]^+$ (Fig. 4.26b) is also observed with sufficiently satisfactory high resolution (within 5 ppm) to be unambiguous, HR MS found 224.0630, $C_{10}H_5D_3Na^{16}O^{18}O_3$ requires 224.0630.

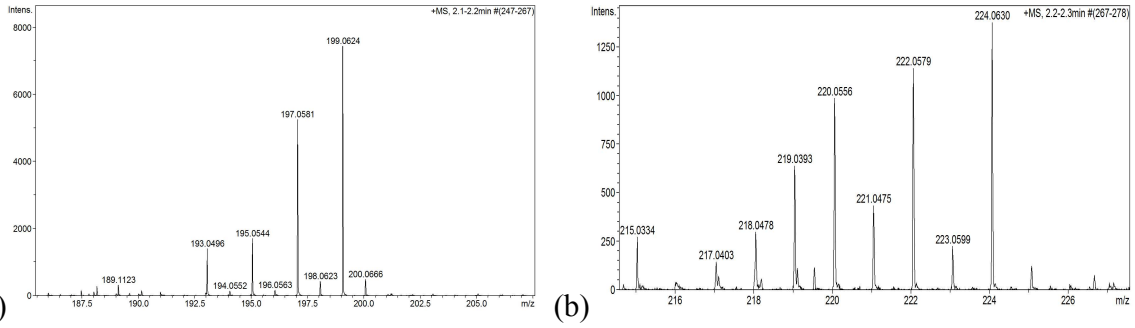


Fig. 4.26. (a) Mass spectrum of scopoletin $[M + H]^+$ after allowing PPD in 20% $^{18}O_2$; (b) Mass spectrum of scopoletin $[M + Na]^+$ after feeding with cinnamic-d_7 acid in 20% $^{18}O_2$.

Summary

We conclude from these results that the accumulation of hydroxycoumarins in cassava roots during PPD makes cassava a good model for studying the biosynthesis of scopoletin and scopolin using various stable isotopic intermediates along the biosynthetic pathway. The major pathway for the biosynthesis of scopoletin and scopolin in cassava roots during PPD was shown to be via *p*-coumaric, caffeic and then ferulic acids which is *o*-hydroxylated, isomerised and lactonised into scopoletin and then glucosylated into scopolin (pathway 3, Fig. 4.22). The experimental evidence given above shows that hydroxycoumarin biosynthesis in cassava during PPD occurs via *o*-hydroxylation and not via a spirodienone intermediate. Although we have no evidence for the order for the *o*-hydroxylation and *E*-*Z*-isomerisation steps, the well-known facile closure of δ-lactones makes lactonisation a rapid step once the *Z*-geometry is established in the presence of a 2'-hydroxyl group. The biosynthesis of esculetin and esculin is via 2',4'-dihydroxycinnamic acid then umbelliferone. Methylation of esculetin to scopoletin occurs only to a small extent. Our phytochemical analysis of secondary metabolite biosynthesis in cassava, both fresh and deteriorated roots, led to the isolation of isotopically labelled scopolin which is highly accumulated in the roots during PPD. As this is the glycoside of scopoletin, this prompted an investigation of hydroxycoumarin glucosyltransferases by screening a PPD-related cDNA library by PCR and nucleic acid hybridization in order to identify the principal glucosyltransferases expressed in the deteriorating cassava root and compare the expression of genes encoding for glucosyltransferases in fresh and deteriorated cassava roots.

Chapter 5

Characterisation of glucosyltransferases in deteriorated cassava roots

5.1. Introduction

Glucosyltransferases

Glucosyltransferases catalyse the transfer of glucose to the aglycone to form a glucoside. In higher plants, secondary metabolites are often converted into their glucoconjugates, which then accumulate in the vacuole. The roles of glucosylation in plants include: solubilisation of the compound in water; detoxification of harmful metabolites or environmental compounds such as herbicides; and the regulation of the action of functional compounds, such as plant hormones (Bowles et al., 2005). Also, glucosyltransferases regulate the activity of compounds that play important roles in plant defence responses against pathogens such as tobacco mosaic virus (Chong et al., 2002) and potato virus (Matros and Mock, 2004).

Glucosyltransferases in cassava

Six different putative UDP-glucose glucosyltransferase clones were isolated from a cassava cotyledon cDNA library probed with an *Ace I-Bgl II* restriction fragment from a UDP-glucose flavonoid 3-*O*-glucosyltransferase from *Antirrhinum majus*. The heterologous probe contained the coding sequence for a glucosyltransferase consensus signature amino acid sequence, which was also found to be present in the isolated cassava cDNA clones (Hughes and Hughes, 1994) Also, two glucosyltransferases from cassava leaves were separated by chromatographic methods and their kinetic properties determined. These enzymes are involved in the last step of biosynthesis of cyanogenic glucosides: the glucosylation of labile hydroxynitriles by UDP-glucose (Mederacke et al., 1996). In another study, four additional glucosyltransferases were also separated from cassava leaf extract, three of which were able to glucosylate α-hydroxynitriles *in vitro*, one proved to be specific for α-hydroxynitriles and the two others were also able to glucosylate flavonoids and anthocyanidins. The fourth glucosyltransferase did not accept hydroxynitriles but revealed high activity towards coumarins, suggesting that glucosylation of the coumarins and the α-hydroxynitriles is catalyzed *in vivo* by different glucosyltransferases (Mederacke et al., 1995).

Glucosyltransferases in deteriorated cassava roots

In chapter 2, the isolation and identification of β-sitosterol glucoside and the galactoside of diacylglyceride from fresh cassava roots, linamarin, esculin and scopolin from deteriorated cassava roots were described. These hydroxycoumarin glucosides, in particular scopolin accumulate in the roots during PPD, and the last step of their biosynthesis is the addition of a glucose unit to esculetin and scopoletin, catalysed by glucosyltransferases to form their glucosides esculin and scopolin (Fig. 5.1). In order to gain further data on these processes clones for glucosyltransferases were isolated from a cassava PPD-related cDNA library and the expression of the corresponding genes were assayed in both fresh and deteriorated cassava roots.

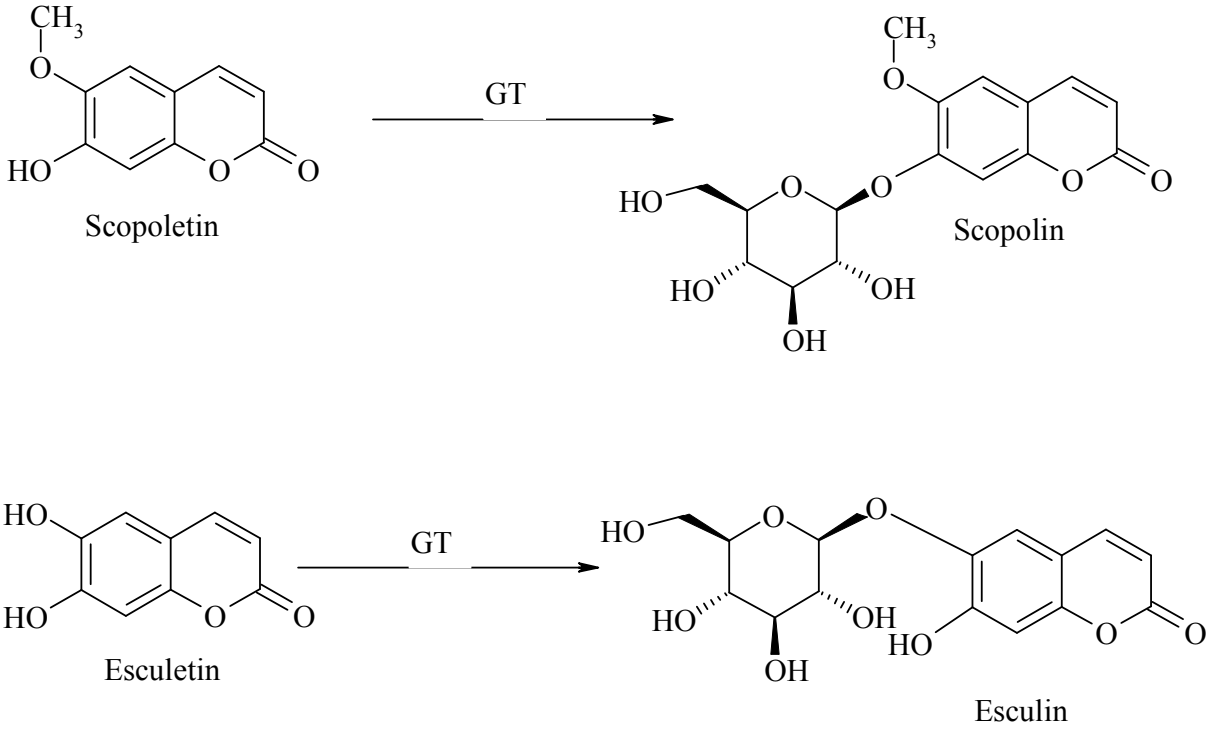


Fig. 5.1. The last steps in the biosynthesis of scopolin and esculin, catalysed by glucosyltransferases (GT).

5.2. Experimental

Plant material

Root tubers (CM2177-2) were harvested after 9 months growth in the tropical glass house at the University of Bath as given above in Chapter 2. The cultivar CM2177-2 shows medium to high susceptibility to PPD, according to CIAT.

cDNA library

Total RNA was extracted and mRNA purified from a range of time-points over a deterioration time course from roots of cassava cultivar CM2177-2. Purified mRNA from time-points 0, 6 and 12 h after harvest were pooled as was mRNA from 24, 48 and 96 h and used to generate "early" and "late" PPD cDNA libraries in Lambda Zap II (Stratagene). These libraries was prepared by Dr Kim Reilly at the University of Bath and should contain cDNA copies of all the genes expressed in the mature root at harvesting and during the first 4 days post-harvest (Reilly et al., 2007).

Bacterial strains

E. coli Top 10

F^- *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*araleu*) 7697 *gal*U *gal*K *rps*L (Str^R) *end*A1 *nup*G. A recombination deficient strain was used for plating and growth of *pCR® 2.2-TOPO®* plasmid. The ϕ80*lac*ZΔM15 allows α complementation with the amino terminus of β-galactosidase encoded in pCR® 2.2-TOPO® plasmid vector.

E. coli XL1-Blue MRF'

Δ(*mcrA*)*183* Δ(*mcrCB-hsdSMR-mrr*)*173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac* [F′ *proAB* $lacI^qZ\Delta M15$ Tn*10* (Tet^r)]. *E. coli XL1-Blue MRF'* is supplied with the Uni-ZAP XR vector kit and is useful for screening the amplified library because the amplified library grows very efficiently on the XL1-Blue MRF′ strain. The F' episome is required for blue/white colour selection. Also it expresses the genes forming the F' pili found on the surface of the bacteria. Without pili formation, filamentous phage (i.e., M13 or f1)

infection could not occur, as the conversion of a recombinant Uni-ZAP XR clone into a pBluescript phagemid requires superinfection with a filamentous helper phage.

E. coli SOLR™ strain

e14⁻(McrA⁻) Δ(*mcrCB-hsdSMR-mrr*)*171 sbcC recB recJ uvrC umuC::*Tn*5* (Kan^r) *lac gyrA96 relA1 thi-1 endA1* λ^R [F' *proAB lacI*^{q}*ZΔM15*] Su– (nonsuppressing), SOLR cells allows only the excised phagemid to replicate in the host, removing the possibility of co-infection from the ExAssist helper phage, as the ExAssist helper phage cannot replicate in the SOLR strain.

Plasmid and phagemid vectors

pCR® 2.2-TOPO® plasmid vector

A 3.9 Kb cloning vector that conferring ampicillin and kanamycin resistance to host cells. It is suitable for blue/white colour screening due to the presence of LacZ α peptide. This region allows selection of recombinant due to β-galactosidase activity. It has M13 forward and reverse priming sites which used for sequencing the cloned PCR product. This plasmid map is shown (Fig. 5.2).

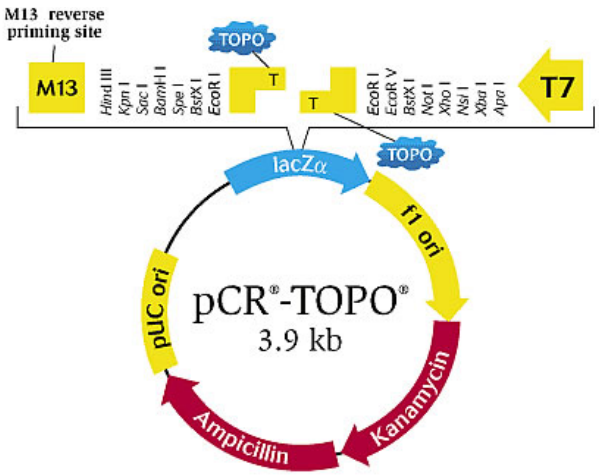


Fig. 5.2. Map shows the features of pCR® 2.2-TOPO® vector. (taken from: http://www.invitrogen.com/site/us/en/home/Products-and-Services/Applications/Cloning/PCR-cloning/PCRC-Misc/General-Subcloning.html).

pBluescript KS (-) phagemid vector

A 3.0 kb standard cloning vector that conferring ampicillin resistance to host cells. And which is suitable for blue/white colour screening. It has M13 forward and reverse priming sites which used for sequencing. A plasmid map is shown (Fig. 5.3). The ExAssist helper phage with SOLR strain is designed to allow efficient excision of the pBluescript phagemid from the Uni-ZAP XR vector.

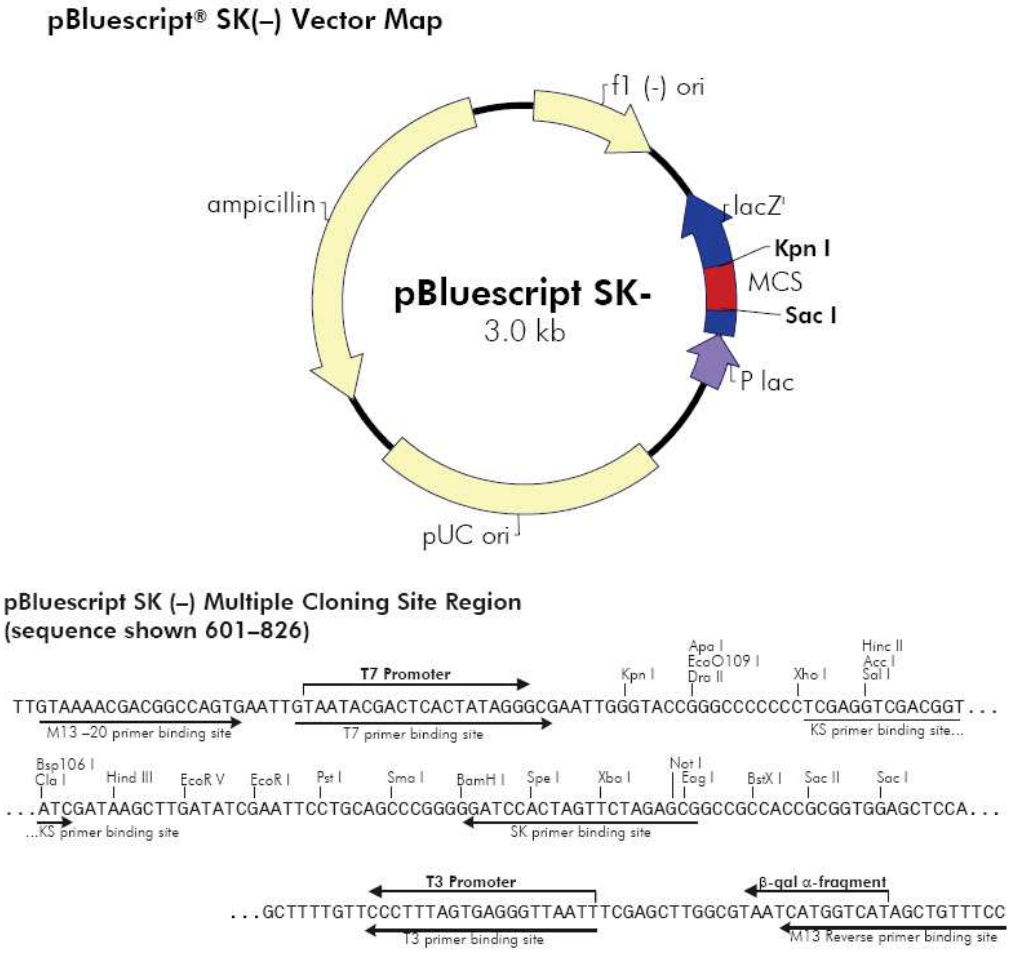


Fig. 5.3 . Map shows the features of pBluescript KS (-) vector (taken from: http://www.stratagene.com/manuals/200401.pdf)

Media for bacterial growth

LB broth

25 g of Luria Broth was dissolved in 800 ml of MilliQ water. MilliQ water was added to make up to 1 l and autoclaved. All autoclaving was performed under these conditions: 10^5 Nm^{-2} for 20 min.

LB broth with supplements

5 ml of sterile 1 M $MgSO_4$ and 5 ml of 20 % (W/V) maltose were added to 500 ml of autoclaved LB broth.

LB agar

40 g of Luria Agar was dissolved in 800 ml of MilliQ water. MilliQ water was added to make up to 1 l and autoclaved.

LB agar plates with supplements

5 ml of sterile 1 M $MgSO_4$ and 5 ml of 20 % (W/V) maltose were added to 500 ml of autoclaved molten LB agar before pouring it into the plates.

NZY top agar

	Per litre
NaCl	5 g
$MgSO_4.7H_2O$	2 g
Yeast extract	5 g
Casein hydrolysate	10 g

These were dissolved in 800 ml of MilliQ water. The pH was adjusted to 7.5 using NaOH, 7 g (0.7 %) of agarose was added and MilliQ water was added to make up to 1l prior to autoclaving.

Reagents and solutions

Reagents and solutions were prepared by reference to Sambrook and Russell (Sambrook and Russell, 2001).

Ampicillin

50 mg/ml ampicillin (sodium salt) dissolved in MilliQ water. Sterilised by filtration. Stored at -20^{o}C.

X-Gal solution

2% (w/v) 5-bromo-4-chloro-3-indoyl-β-D-galactoside made up in dimethylformamide and stored at -20 ^{o}C.

10 mM $MgSO_4$

This was made by dilution of 1 M solution prepared by dissolving 24.65 g of magnesium sulfate heptahydrate in 75 ml of MilliQ water. The volume was made up to 100 ml with MilliQ water and autoclaved.

20 % Maltose

10 g of maltose was dissolved in 50 ml of MilliQ water and sterilised by filtration through 0.22 µm and stored at 4 °C.

0.5 M EDTA (pH 8.0)

186.1 g of disodium ethylene diamine tetra acetic acid dihydrate was dissolved in 800 ml of MilliQ water. The pH was adjusted to 8.0 using NaOH pellets. MilliQ water was added to make up to 1l and autoclaved.

100 mM Tris HCl (pH 7.5)

12.1 g of tris base (tris(hydroxymethyl) methylamine) was dissolved in 800 ml of MilliQ water. The pH was adjusted to 7.5 using conc. HCl, MilliQ water was added to make up to 1 l and autoclaved.

2 % Gelatine

2 g of gelatine was dissolved in 100 ml of MilliQ water and autoclaved.

20 % SDS

100 g of sodium dodecyl sulphate was dissolved in 400 ml of MilliQ water. MilliQ water was added to make up to 500 ml.

SM Buffer

	Per litre	final concentration
NaCl	5.8 g	0.1 M
$MgSO_4.7H_2O$	2.0 g	0.01 M
Tris base	6.05 g	0.05 M
2% gelatine	5 ml	0.01%

These were dissolved in 800 ml of MilliQ water. The pH was adjusted to 7.5 with conc. HCl and the volume was made up to 1 l with MilliQ water and autoclaved.

1 M Sodium phosphate buffer (pH 7.2)

A stock solution of 1 M sodium dihydrogen phosphate and 1 M disodium monohydrogen phosphate was prepared and autoclaved. To 68.4 ml of 1 M disodium monohydrogen phosphate 31.6 ml of 1 M sodium dihydrogen phosphate was added till pH 7.2 was reached. The solution was sterilised by autoclaving.

Electrophoresis buffers

Tris-Borate EDTA (TBE)

5X stock solution

	Per litre	final concentration
Tris base	54 g	0.45 M
Boric acid	27.5 g	0.45 M
0.5 M EDTA (pH 8.0)	20 ml	0.01 M

Made up with MilliQ water to 1 l then autoclaved.

Tris-Acetate EDTA (TAE)

50X stock solution

	Per litre	final concentration
Tris base	242.28 g	2 M
Glacial acetic acid	57.1 ml	1 M
0.5 M EDTA (pH 8.0)	100 ml	0.05 M

Made up with MilliQ water to 1 l then autoclaved.

Extraction Buffer (RNA extraction)

	Per litre	final concentration
1 M Tris HCl (pH 7.5)	100 ml	100 mM
5 M NaCl	20 ml	100 mM
500 mM EDTA	50 ml	25 mM
SDS	10 g	1 % (w/v)
PVP K30	20 g	2 %(w/v)

Glassware was treated at 180 °C for about 4 h, the solution was treated with 0.1 % (v/v) DEPC for at least 1 h in a fume cupboard and then autoclaved.

DEPC treated water/ solution

0.1 % (v/v) of DEPC was added to the water solution in a fume cupboard. The solution was mixed and allowed to stand for at least 1 h and then autoclaved.

Preparation of replica filters for screening the cDNA library

Denaturation solution

	Per litre	final concentration
NaCl	87.66 g	1.5 M
NaOH	20 g	0.5 M

These were dissolved in 800 ml of MilliQ water. The volume was made up to 1 l with MilliQ water.

Neutralisation solution

	Per litre	final concentration
NaCl	87.66 g	1.5 M
Tris base	60.5 g	1.5 M

These were dissolved in 800 ml of MilliQ water. The pH was adjusted to 7.5 with conc. HCl. The volume was made up to 1 l with MilliQ water.

Nucleic acid transfer buffer (20X SSC)

	per litre	final concentration
Tri-sodium citrate	88.23 g	0.3 M
NaCl	175.32 g	3.0 M

These were dissolved in 800 ml of MilliQ water. The pH was checked to 7.5 with conc. HCl. The volume was made up to 1 l with MilliQ water.

Hybridisation with radiolabelled probes

Pre-hybridisation / hybridisation solution

	per 500 ml	final concentration
1 M sodium phosphate buffer (pH 7.2)	250 ml	500 mM
0.5 M EDTA (pH 8.0)	10 ml	10 mM
20% SDS	175 ml	7%

Made up with MilliQ water to 500 ml.

Low stringency wash solution (Equivalent to 1.7X SSC)

	per 500 ml	final concentration
1 M sodium phosphate buffer (pH 7.2)	62.5 ml	125 mM
20% SDS	2.5 ml	0.1%

Made up with MilliQ water to 500 ml.

Medium stringency wash solution (Equivalent to 0.8X SSC)

	per 500 ml	final concentration
1 M sodium phosphate buffer (pH 7.2)	31.5 ml	62.5 mM
20% SDS	2.5 ml	0.1%

Made up with MilliQ water to 500 ml.

High stringency wash solution (Equivalent to 0.2X SSC)

	per 500 ml	final concentration
1 M sodium phosphate buffer (pH 7.2)	7.5 ml	15 mM
20% SDS	2.5 ml	0.1%

Made up with MilliQ water to 500 ml.

Experimental methods

Sequences analysis

DNA or amino acid sequences were retrieved from the NCBI (The National Center for Biotechnology Information) databases (http://www.ncbi.nlm.nih.gov). Multiple sequence alignments (using CLUSTAL X of telnet genome GCG package) (Devereux et al., 1984) were performed to generate Neighbour-Joining phylogenetic trees using 1000 bootstrap trials. The alignments were coloured using GENEDOC (version 2.7) (http://www.psc.edu/biomed/genedoc). TREEVIEW (version 1.6.6) was used to produce the phylogenetic trees (Page, 1996) from the matrices generated from CLUSTAL X.

PCR reactions

Each PCR sample was performed using *Taq* DNA polymerase with Standard *Taq* Buffer (Biolabs, # M0273) in a reaction containing the following: forward primer (0.8 μM), reverse primer (0.8 μM), ~100 ng template DNA, dNTPs (0.25 mM), 2.5 U *Taq* DNA polymerase (Biolabs, # M0273), 1 × Standard *Taq* reaction buffer and SDW 50 μl. Samples were cycled in a PTC-200 DNA Engine (MJ Research, Inc.); using the

following conditions unless otherwise stated: 94 ºC for 3 min, [94 ºC for 1 min, 45-55 ºC for 1 min, 72 ºC for 2 min 35 cycles] and 72 ºC for 10 min.

The PCR product was gel isolated using the QIAEX®II Gel Extraction Kit (Qiagen, # 20021) according to the manufacture's specifications, cloned into PCR®2.1-TOPO® vector and then transformed to One Shot® Top10 Chemically competent *E. coli* cells using TOPO TA Cloning® Kit (Invitrogen, # K4500-01). To allow for blue/white colour screening of transformed colonies 80 µl of X-Gal (2% w/v) was spread over the surface of LB agar plates supplemented with ampicillin (50 µg/ml final concentration) using a sterile spreader. Plates were allowed to dry for 1-2 h. Transformed cells in 250 µl of LB broth were then spread in 20, 50, 150 µl aliquots onto the previously prepared LB agar plates and residual liquid was allowed to absorb before incubating inverted plates at 37 ^{o}C overnight. A single white colony (2-3mm diameter) was picked from a freshly grown plate and used to inoculate 2 ml of LB broth supplemented with ampicillin (50µg/ml final concentration). The culture was grown overnight at 37 ^{o}C and 150 rpm. The plasmid DNA was purified from the grown culture using QIAprep® Spin Miniprep Kit (250) (Qiagen, # 27106).

Electrophoresis of DNA on Agarose gels

Agarose gels were prepared in 1X TBE buffer or 1X TAE buffer. The buffer was added to agarose (the final concentration was 1%) and melted by several 30 second pulses in a microwave oven. Ethidium bromide was added after the molten gel had been cooled to about 55 °C at a final concentration of 0.5 µg/ml. Gel slabs were poured out to a thickness of 0.5-1 cm with a gel comb in place to give sample wells of appropriate dimensions. Running buffer was the same buffer as that used for the gel preparation. DNA samples were diluted if necessary with MilliQ water and 6X gel loading buffer was added to a final concentration of 1X per sample. Ladders (1 kb or 100 bp) were used as molecular weight markers. Electrophoresis was carried out 2-8 V/cm for 1-5 h. DNA bands were visualised and documented under UV light using an UVP white /UV transilluminator and digital graphic printer.

DNA sequencing and identification

The DNA was sequenced, each sequencing reaction contained 4 µl BigDye®Terminator v 3.1 (Applied Biosystems, # 4337455), 3 µl primer (1 µm), 500 ng DNA and SDW to 10 µl. Reaction was performed in a PTC-200 DNA Engine (MJ Research, Inc.) using 25 cycles of: 96 ºC for 30 s, 50 ºC for 15 s and 60 ºC for 4 min. Primers used for sequencing are M13 forward and reverse primers:

M13F: $^{5'}$GTAAAACGACGGCCAG$^{3'}$

M13R: $^{5'}$CAGGAAACAGCTATGAC$^{3'}$

DNA was precipitated by isopropanol precipitation. DNA sequence was resolved by Lark Technologies Inc. (Essex, UK). Returned trace files were edited using trev 1.6 through GCG software version 11.1 (Accelrys Inc.) and BLAST searches were done using NCBI (http://www.ncbi.nlm.nih.gov/BLAST/).

Isopropanol precipitation

Following the sequencing reaction, the reaction mix was transferred to a 0.5 ml microfuge tube then 10 µl of SDW, 30 µl of isopropanol were added, the content was mixed by flicking the tube, the tube was left to stand for 10 min then centrifuged (150 rcf) for 30 min. The isopropanol was removed carefully using a vacuum aspirator. The DNA pellet was washed with 100 µl of 75 % isopropanol then centrifuged for 10 min; the 75 % isopropanol was removed carefully using vacuum aspirator and the pellet was dried in air for 30 min.

Restriction digestion

DNA solutions of known concentration were mixed with restriction enzymes in appropriate buffer according to the specifications of the manufacturer (Kramel Biotechnologies, New England Biolabs, Bioline or Advanced Biotechnologies), and the reactions were incubated at 37^{o}C for 1 to 2 h. Care was taken to ensure that the volume of restriction enzyme did not exceed 1/10 the final reaction volume, in order to avoid inhibition of enzyme activity by glycerol in which commercial enzyme preparations are stored. The reaction was stopped by transferring the tubes to an ice bath.

cDNA library screening

Preparation of plating cells

A suspension of plating cells (XL1-Blue MRF′ strain) was prepared by inoculating a single colony from fresh plate in 50 ml of LB broth with supplements (0.2% (W/V) maltose and 10 mM $MgSO_4$) and incubating overnight with shaking at 30 ºC. The culture was centrifuged at 1400 rcf for 10 min then the cells were gently resuspended in 10 ml of 10 mM $MgSO_4$. Before use cells were diluted to an OD_{600} = 0.5.

Infection of plating cells with phage

Lambda phage particles containing the cDNA library of deteriorated cassava root were diluted in SM buffer to give the desired number of pfu/ml. For small (90 mm diameter) LB agar plates with supplements 250 µl of plating cell suspension and 1 µl of diluted lambda (150 pfu) were gently mixed in a Falcon tube and incubated at 37 ºC for 20 min to allow the phage to adsorb to the cells. For large (140 mm diameter) plates 600 µl of plating cell suspension and 1 µl (250 pfu) of diluted lambda were used. 4-8 ml of molten NZY top agar (~45 ºC) was added to each tube and gently mixed before pouring on to the pre-warmed LB agar with supplements plate. The plates were allowed to set and slightly air dry for about 10 min. They were incubated inverted for 10-16 h.

Preparation of replica filters for screening the cDNA library

Plates for screening were left at least 1 h at 4 ºC. Positively charged Nylon transfer membrane (Amersham, # RPN132B) was placed over the surface of the plate for 1 min for the 1st membrane and for 3 min for the replicate membrane. The membrane was marked in three or more asymmetric positions by stabbing through it and into the agar beneath with an 18-gauge needle. The membrane was removed using blunt ended forceps and placed plaque side up on a piece of Whatman 3MM paper soaked in 10% SDS for 3 min. It was transferred to another piece of Whatman 3MM paper soaked in denaturation solution for 5 min. The filter was subsequently transferred to two trays each containing Whatman 3MM paper soaked in neutralizing solution each for 3 min. The membrane was washed in 2×SSC solution and air–dried for 30 min.

DNA fixation to nylon membranes

The membranes were put between two Whatman 3MM papers and baked for 2 h at 80 ºC. Membranes were immediately used for hybridisation or wrapped in Saran film and stored at –20 ºC.

Hybridisation with radiolabelled probes

Membranes from plaque lifts were placed in a plastic box and prehybridisation solution (100 ml) was added. The membranes were prehybridised for at least 1 h at 62 ºC in an Hybaid hybridisation oven. DNA (50 ng) was labelled with (α- ^{32}P)dCTP (Amersham, #AA0005-250UCL) using Random Primers DNA Labeling System kit (Invitrogen, # 18187_013) according to the manufacturer's specifications. Labelled probe was denatured by heating at 100 ºC for 5 min and snap cooling on ice for 2 min. Prehybridisation solution was poured off and replaced with 50 ml of fresh hybridization solution. Labelled probe was added to the solution and hybridization was allowed to proceed for 17 h at 62 ºC. The membranes were rinsed with low stringency solution (2 x 100 ml) at 62 ºC for 15 min, then washed with medium stringency solution (100 ml) at 62 ºC for 15 min. The excess of liquid on the membranes was then removed with the edge of a small plastic box, the membranes were wrapped in Saran film and placed in an autoradiography cassette with intensifying screens. The membranes were covered with a sheet of X-ray film (Kodak) and autoradiograph cassette stored at –70 ºC. Autoradiographic films were developed using an X-omat processor.

Single clone excision

Preparation of lambda-phage stock

The plaques of interest were cored out from the plate using a Pasteur pipette and put into a micro centrifuge tube containing 500 µl of SM buffer and 20 µl of chloroform and then vortexed. The tubes were incubated overnight at 4 ºC.

Preparation of plating cells

Single colony from fresh plates of XL1-Blue MRF´ and SORL strain (Stratagene, # 200400) were inoculated and incubated separately into 50 ml of LB broth with supplements and incubated overnight with shaking at 30 ºC. The culture was centrifuged at 1400 rcf for 10 min then the cells were gently resuspended in 10 ml of 10 mM $MgSO_4$. Before use cells were diluted to an $OD_{600} = 1$.

Preparation of the excised phagemid

In a Falcon tube, 200 µl of XL1-Blue MRF´, 250 µl of phage stock and 1 µl of ExAssist helper phage were added. The tube was incubated at 37 ºC for 20 min then 3 ml of LB broth with supplements was added. The tube was incubated for 3 h at 37 ºC with shaking. The tube was heated at 65-70 ºC for 20 min then centrifuged at 1400 rcf for 15 min. The supernatant containing the excised pBluescript phagemid packaged as filamentous phage particles was decanted into a sterile Falcon tube and stored at 4 ºC.

Preparation of plates of the excised phagemid

200 µl of SORL was added to two centrifuge tubes and 100 µl of the phage supernatant was added to one tube and 10 µl of the phage supernatant was added to the other tube. The tubes were incubated at 37 ºC for 20 min. 100 µl and 50 µl from each tube were plated on LB-ampicillin agar plates (100 µg/ml). The plates were incubated overnight at 37 ºC. The phagemid DNA was then purified using QIAprep® Spin Miniprep Kit (250) (Qiagen, # 27106). DNA was sequenced and identified.

Cassava root RNA extraction

Root tissue for RNA extraction was grated, wrapped in aluminium foil and immediately frozen in liquid nitrogen. Where the sample was not processed immediately, it was stored at -70 °C until required. Grated tissue was homogenised to a fine powder using a porcelain pestle and mortar previously chilled at -70 °C, aliquots of liquid nitrogen were added continuously during grinding to prevent thawing and consequent degradation by RNases in the sample. 0.3 g of powdered material was transferred to a 1.5 ml RNase free tube that had been chilled by immersion in liquid

nitrogen. 800 µl of extraction buffer (pre-warmed at 60 °C), 20 µl of mercaptoethanol and 800 µl of chloroform: isoamyl alcohol (24:1) were added to the tube, which was vortexed and then incubated at room temperature for 5 min. The mixture was centrifuged at 150 rcf for 2 min at 4 °C. The upper layer was transferred to a 1.5 ml tube and the same volume was added of chloroform: isoamyl alcohol (24: 1), the tube vortexed and then centrifuged for 2 min at 4 °C. The upper layer was transferred to a 1.5 ml tube, 0.5 volume of Trizol® Reagent (Tri Reagent, Sigma, T9424) was added, the tube shaken by hand for 15 s, incubated at 20 °C for 5 min, then 0.2 ml of chloroform per each 1 ml of Trizol® Reagent was added, the tube was shaken by hand for 15 s, and then incubated at 20 °C for 5 min. The mixture was centrifuged for 2 min at 4 °C. The upper layer was transferred to a 1.5 ml tube and 0.5 ml of isopropanol per each 1ml of Trizol® Reagent was added to precipitate RNA. The mixture was mixed gently by hand and incubated at 20 °C for 10-30 min and then centrifuged for 5 min at 4 °C. The supernatant was discarded and the pellet was washed with 70% ethanol then centrifuged for 2 min at 4 °C. The supernatant was discarded and the pellets dried at 20 °C for 10 min and then dissolved in 50 µl treated water. To assess the quality and quantity of RNA, aliquots were run on a 2% TAE gel and the absorbance measured at 260 and 280 nm.

DNA removal

RNA was purified from any remaining DNA using DNA-free™ Kit according to the specifications of the manufacturer (Ambion).

RT-PCR reaction

Each RT-PCR sample was performed using an Access RT-PCR System Kit (Promega, # A1260) in a reaction containing the following: 0.25 µl forward primer (1 µM), 0.25 µl reverse primer (1 µM), 2 µl template RNA (20 ng), 0.5 µl dNTPs (0.2 mM), 1 µl of 25 mM $MgSO_4$ (1 mM), 5 µl AMV / *Tfl* reaction buffer (1X), 0.5 µl AMV reverse transcriptase (0.1u / µl), 0.5 µ l *Tfl* DNA polymerase(0.1u / µl) and nuclease free water to 25 µl. Samples were cycled in a PTC-200 DNA Engine (MJ Research, Inc.) using the following conditions: 45 ºC for 45 min, 94 ºC for 2 min, [94 ºC for 30 s, 55-60

ºC for 1 min, 68 ºC for 2 min 40 cycles] and 68 ºC for 7 min. Agarose gel electrophoresis was carried out on the product. DNA bands were visualised and documented under UV light.

5.3. Results and discussion

Characterisation of some genes and enzymes known to be involved in the biosynthesis of scopoletin and scopolin.

A search for the enzymes known to be involved in the three proposed pathways for the biosynthesis of scopoletin and scopolin (Fig. 5.4) revealed that few of them or their genes have been identified and fully characterised. Indeed some have only been identified by the help of tracer feeding experiments using radio labelled intermediates. In order to assist the process of characterising these pathways, in the following, some of the available data on these enzymes and their genes are collated and evaluated.

Phenylalanine ammonia lyase (PAL)

Phenylalanine ammonia lyase (PAL, E.C.4.3.1.5) catalyses the first committed step in the phenylpropanoid biosynthetic pathway, which is the non-oxidative deamination of L-phenylalanine to *E*-cinnamic acid (Fig. 5.4). PAL was isolated from *Hordeum vulgare* (Koukol and Conn, 1961) and subsequently was identified from various plants e.g. *Arabidopsis* and parsley (Cramer et al., 1989; Logemann et al., 1995; Cochrane et al., 2004; MacDonald and D'Cunha, 2007).

PAL (GenBank accession numbers AY036011, AF383152, AF383150 and AF078690) has been identified in cassava. There was a progressive increase of PAL activity in cassava suspension cells after yeast elicitor treatment (Gómez-Vásquez et al., 2004). PAL enzyme activity was significantly higher during the resistant interaction against bacterial pathogen (*X. axonopodis* pv. *manihotis,* which causes cassava bacterial blight). Also PAL activity increased in cassava roots pot-harvesting (Tanaka et al., 1983; Pereira et al., 1999).

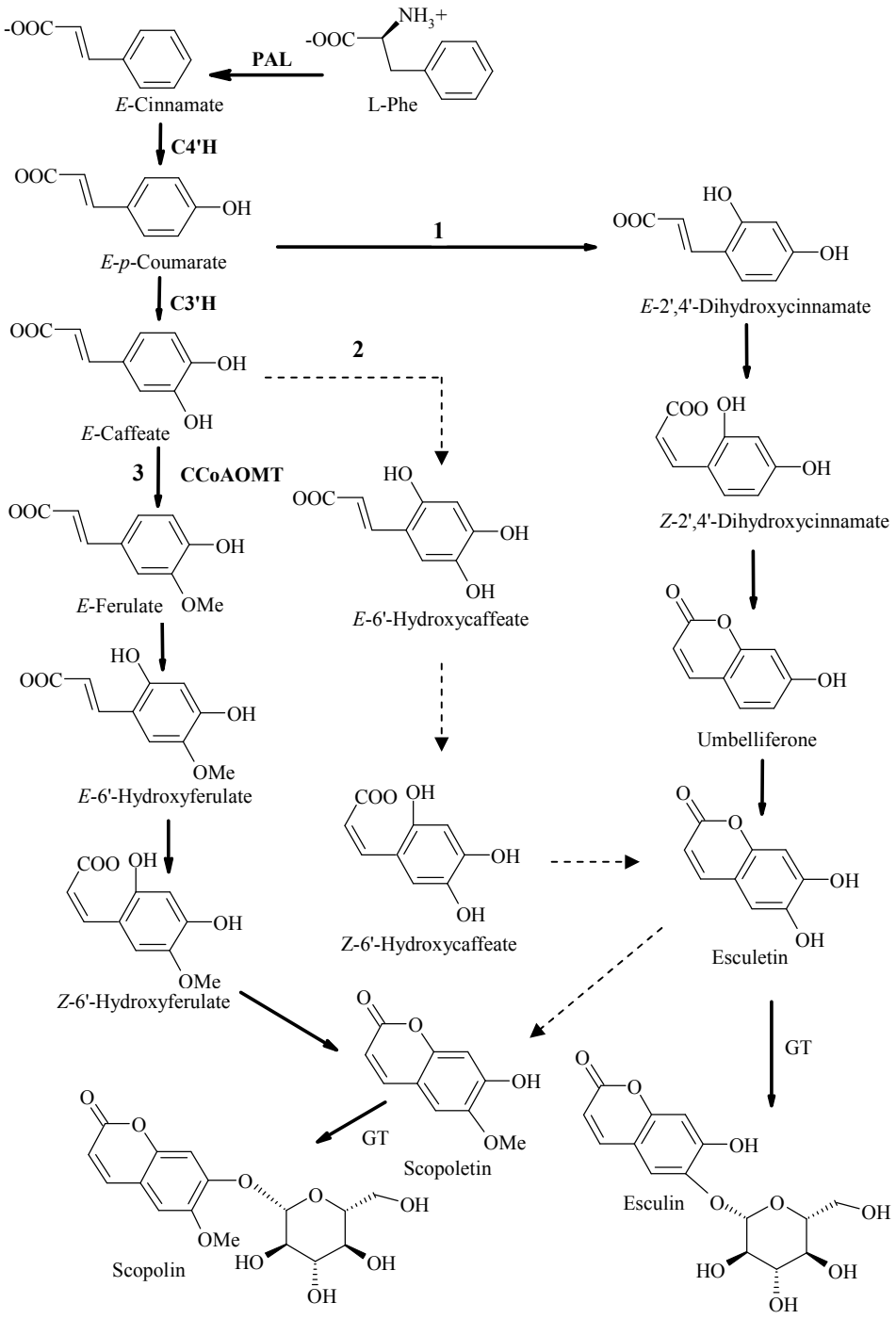


Fig. 5.4. Proposed metabolic pathways of the biosynthesis of scopoletin in plants showing genes identified in different plants. Abbreviations: cinnamate-4'-hydroxylase (C4'H), *p*-coumaroyl shikimate/quinate 3'-hydroxylase (C3'H), caffeoyl-CoA *O*-methyltransferase (CCoAOMT) and glucosyltransferase (GT).

Cinnamate-4'-hydroxylase (C4'H, Cyp73s)

This is a cytochrome P450 monoxygenase (**EC** 1.14.13.11) that catalyses cinnamate-4'-hydroxylation (Fig. 5.4). It was one of the first P450s to be characterised in plants (Nair and Vining, 1965; Russell and Conn, 1967) and subsequently it has been studied in a variety of plants (Table 5.1). It is a major flux control for lignification and so has been engineered to alter its expression in order to regulate phenolic metabolism (ex. tobacco and tomato) for diverse purposes, such as increased levels of antioxidant flavonoids for dietary improvements or reduction of lignin for improvements to fibre resources for industrial use (Millar et al., 2007).

Full-length cDNA sequences encoding higher plant C4'H (Table 5.1) were retrieved from the NCBI databases, aligned and used to generate a Neighbour-Joining phylogenetic tree (Fig. 5.5). The trees were designed to determine the relationship of these to each other. The phylogenetic tree for C4'H rooted to NM_001051180 from *Oryza sativa*. The tree shows separate clusters of monocotyledons and dicotyledons genes. This could be due to gene duplication occurring after the separation of plants into monocotyledons and dicotyledons. The tree shows high similarity of genes in the same plant species for example all *Brassica* genes are in one cluster while all *Arabidopsis* genes except NM_120023 are in another cluster, which suggests that some gene duplication events are even more recent.

Table 5.1

Cinnamate-4'-hydroxylase (C4'H)

Accession number	Source	Reference
AY219918	*Ammi majus*	Unpublished
NM_120023	*Arabidopsis thaliana*	Unpublished
NM_128601	*Arabidopsis thaliana*	Unpublished
BT008875	*Arabidopsis thaliana*	Unpublished
U37235	*Arabidopsis thaliana*	Unpublished
U71081	*Arabidopsis thaliana*	(BellLelong et al., 1997)
U71080	*Arabidopsis thaliana*	(BellLelong et al., 1997)
DQ485132	*Brassica napus* (rape)	(Chen et al., 2007)
DQ485130	*Brassica napus*	(Chen et al., 2007)
DQ485129	*Brassica napus*	(Chen et al., 2007)
Z32563	*Catharanthus roseus*)	Unpublished
AY641731	*Camellia sinensis* (tea)	Unpublished
AF378333	*Citrus x paradise* (grapefruit)	Unpublished
AF255014	*Citrus sinensis* (sweet orange)	(Betz et al., 2001)
AF255013	*Citrus sinensis*	(Betz et al., 2001)
Z17369	*Helianthus tuberosus*	(Teutsch et al., 1993)
DQ075002	*Malus x domestica* (apple)	Unpublished
AF097664	*Mesembryanthemum crystallinum* (common iceplant)	Unpublished
NM_001061725	*Oryza sativa* (Japanese rice)	(Matsumoto et al., 2005)
NM_001053349	*Oryza sativa*	(Matsumoto et al., 2005)
NM_001053354	*Oryza sativa*	(Ohyanagi et al., 2006)
NM_001051180	*Oryza sativa*	(Ohyanagi et al., 2006)
AB207105	*Oryza sativa*	(Yang et al., 2005)
DQ211885	*Parthenocissus henryana*	Unpublished
L38898	*Petroselinum crispum* (parsley)	(Logemann et al., 1995)
Y09449	*Phaseolus vulgaris* (kidney bean)	(Nedelkina et al., 1999)
Y09448	*Phaseolus vulgaris*	(Nedelkina et al., 1999)
Y09447	*Phaseolus vulgaris*	(Nedelkina et al., 1999)
AY764926	*Pinus taeda* (loblolly pine)	(Brown et al., 2004)
AF096998	*Pinus taeda*	(Anterola et al., 2002)
DQ522293	*Populus tremuloides*	(Lu et al., 2006)
DQ522292	*Populus tremuloides*	(Lu et al., 2006)
DQ522294	*Populus tremuloides*	(Lu et al., 2006)
U47293	*Populus tremuloides*	Unpublished
AF302495	*Populus trichocarpa X Populus deltoides*	(Ro et al., 2001)
AJ309127	*Ruta graveolens* (rue)	Unpublished
EF377337	*Salvia miltiorrhiza*	Unpublished
L07634	*Vigna radiate* (mung bean)	(Mizutani et al., 1993)

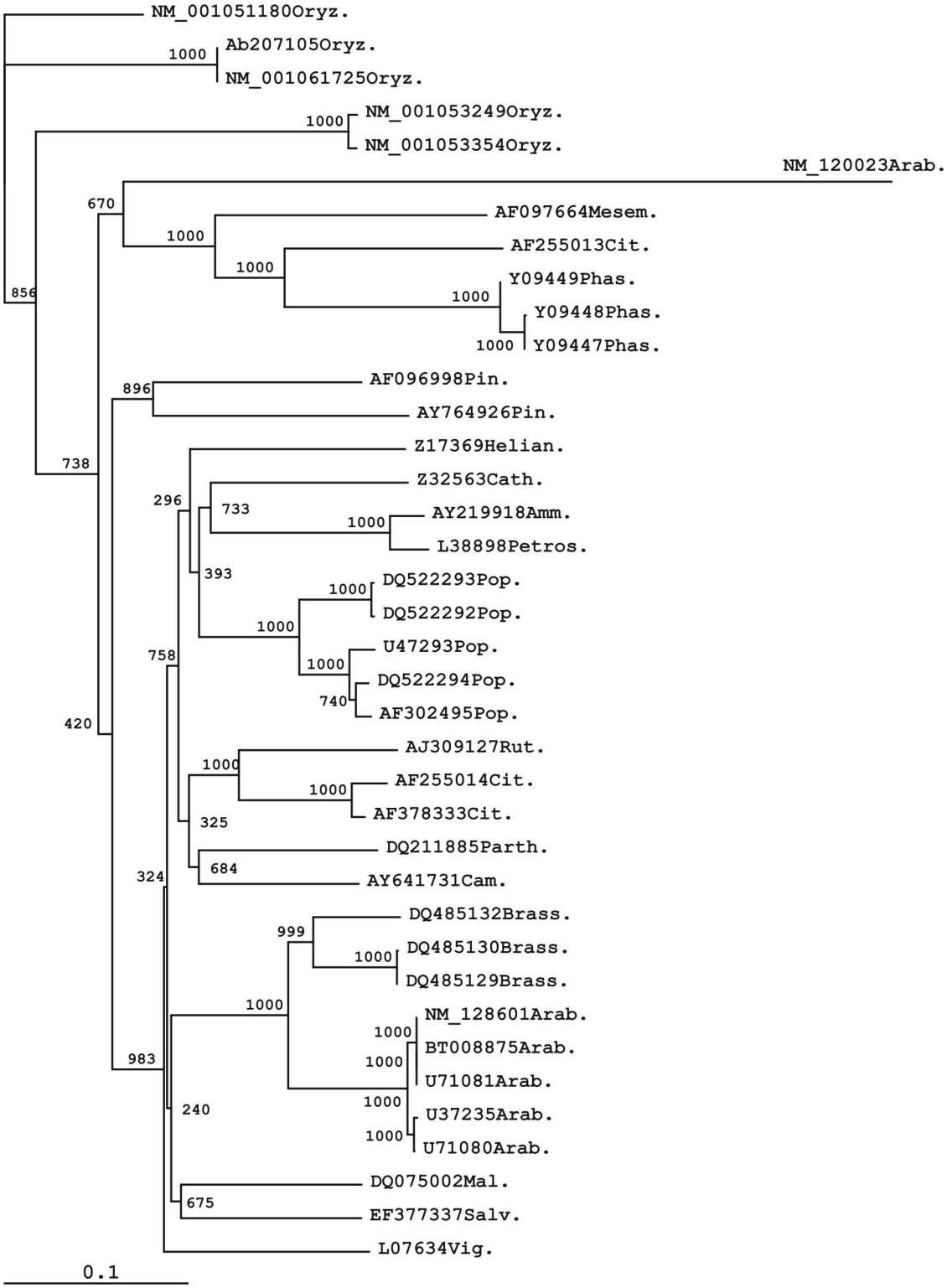


Fig. 5.5. Neighbour-Join bootstrap tree illustrating sequence similarity relationships of DNA sequences of cinnamate-4'-hydroxylases (C4'H) in plants.

p-Coumaroyl shikimate/quinate 3'-hydroxylase (C3'H, CYP98)

This is a cytochrome P450 monoxygenase that catalyses 3'-hydroxylation of *p*-coumarate to produce caffeate (Fig. 5.6) (Ehlting et al., 2006). It was found that the shikimate and quinate esters are very actively hydroxylated in the 3'-position of phenolic rings (Schoch et al., 2001). In *Arabidopsis thaliana,* coumarin (scopoletin and scopolin) content in mutants that carried a T-DNA insertion within the gene encoding CYP98A3, decreased to 3% of that in the wild-type roots. This observation suggested that scopoletin and scopolin biosynthesis in *Arabidopsis thaliana* is strongly dependent on the 3'-hydroxylation of *p*-coumarate units catalyzed by CYP98A3 (Kai et al., 2006). Only a few genes for this enzyme have been identified from plants (Table 5.2).

cDNA sequences encoding C3'H (Table 5.2) were retrieved from the NCBI databases and partial sequences were excluded. Multiple sequence alignment of cDNA sequences of C3'H (Fig. 5.7) showed high similarity to each other although they are from four different plants (*Arabidopsis thaliana*, *Coffea canephora*, *Ocimum basilicum* and *Sesamum indicum*). The presence of conserved regions along the length of the sequences could be used to design degenerate primers to amplify this gene in the cDNA library of deteriorated cassava roots.

Table 5.2

p-coumaroyl shikimate/quinate 3'-hydroxylase (*p*-coumarate 3'-hydroxylase) (C3'H**)**

Accession number	**Source**	**Reference**
NM_180006	*Arabidopsis thaliana*	(Kai et al., 2006)
DQ269127	*Coffea canephora* **(**robusta coffee)	(Mahesh et al., 2007)
DQ269126	*Coffea canephora*	(Mahesh et al., 2007)
AY082611	*Ocimum basilicum* (sweet basil)	(Gang et al., 2002)
AY082612	*Ocimum basilicum*	(Gang et al., 2002)
AY065995	*Sesamum indicum* (sesame)	(Anterola et al., 2002)

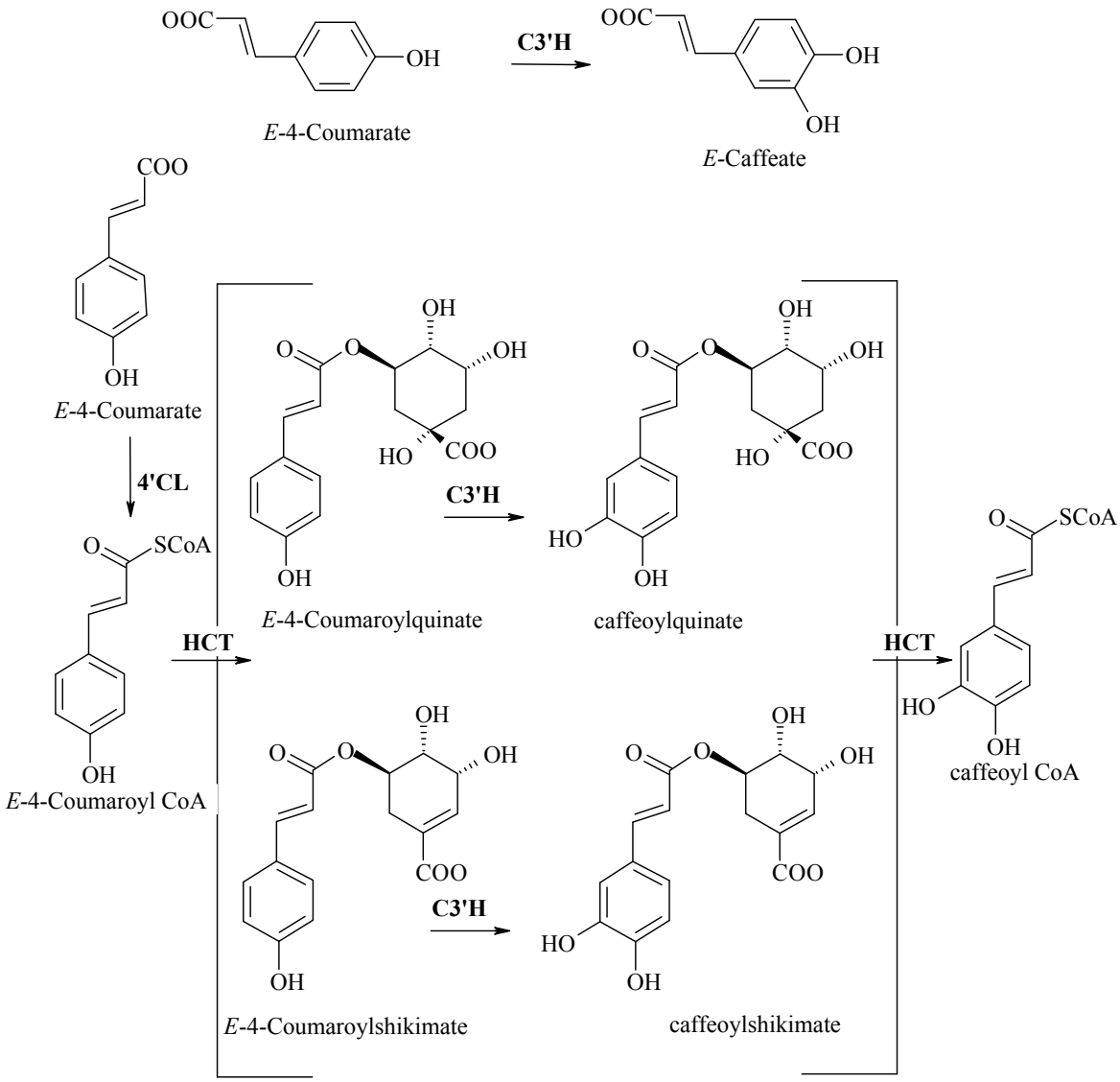
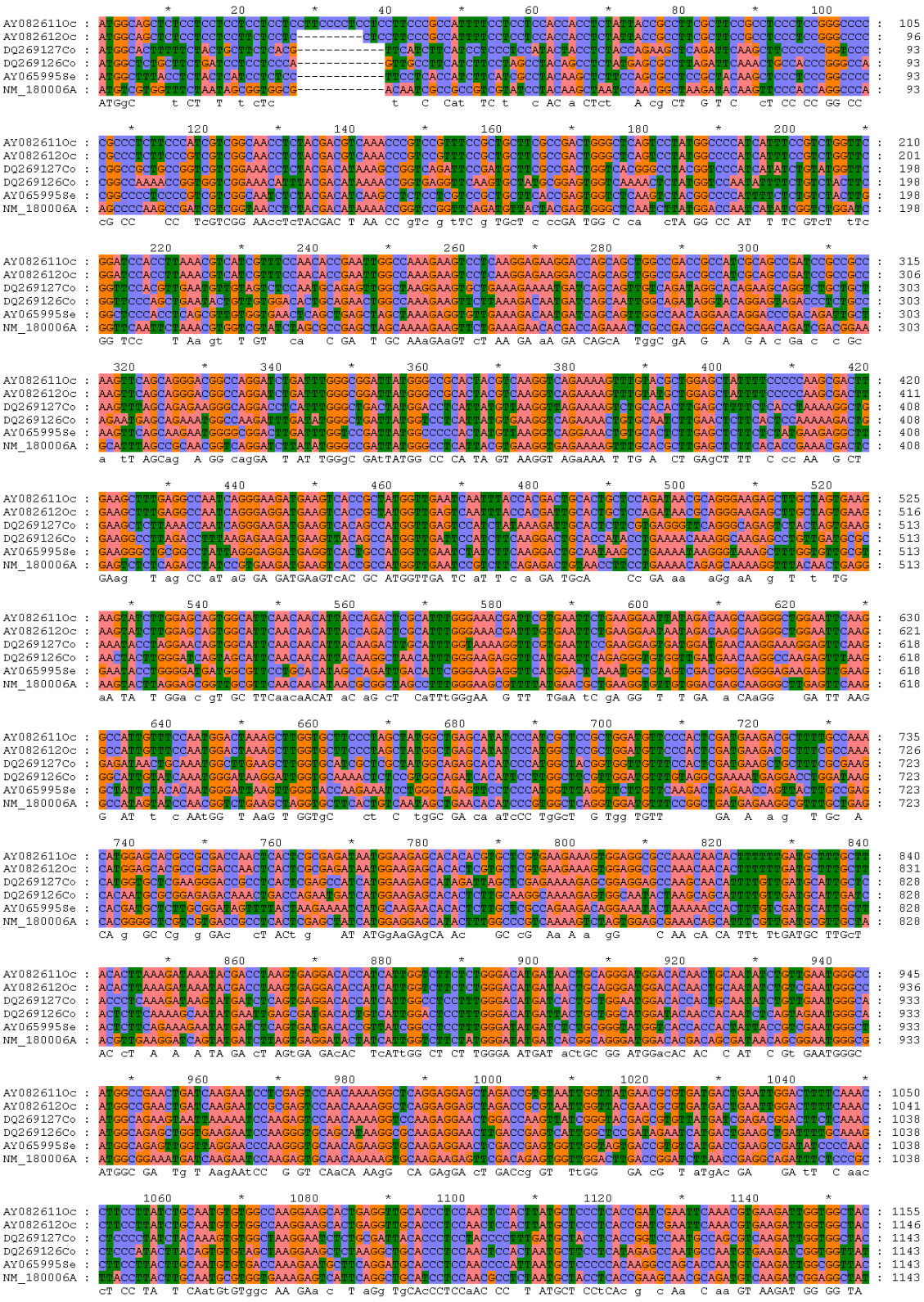


Fig. 5.6. Biosynthesis of caffeate from *p*-coumarate (Kai et al., 2006; Ehlting et al., 2006). Abbreviations: *p*-coumaroylshikimate/quinate-3'-hydroxylase (C3'H), 4'-hydroxycinnamoyl CoA ligase (4'CL), shikimate/quinate hydroxycinnamoyltransferase (HCT).



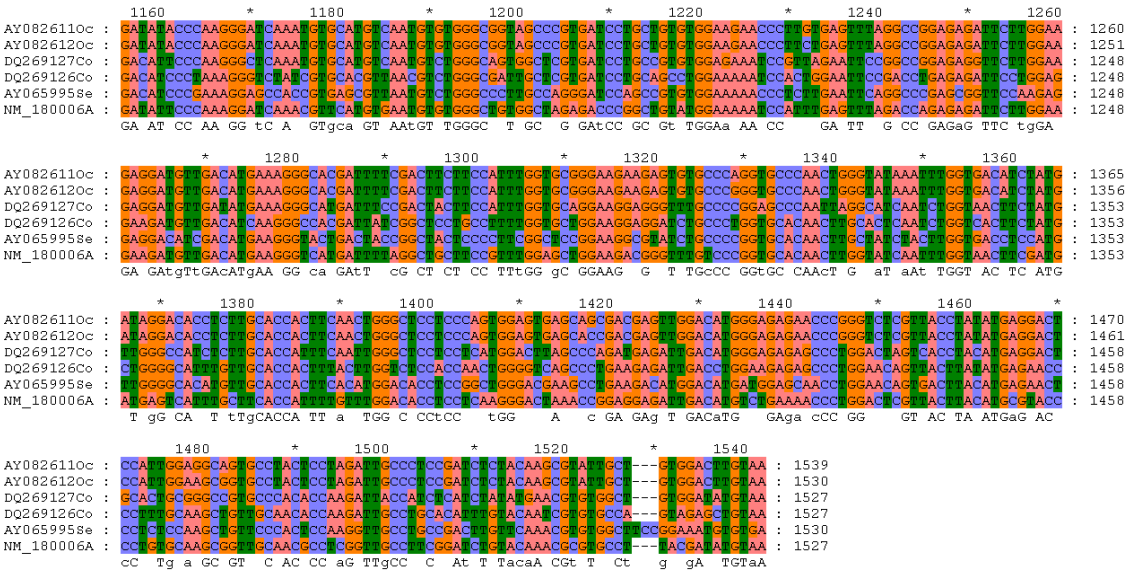


Fig.5.7. Nucleotide alignments of *p*-coumaroyl shikimate/quinate 3'-hydroxylases. Adenine nucleotides (A) are shown in blue, Cytosine (C) in green, Guanine (G) in orange and Thymine (T) in pink.

Caffeoyl-CoA O-methyltransferase (CCoAOMT) and caffeic acid methyltransferase(COMT)

These enzymes catalyse the transfer of a methyl group from S-adenosyl-L-methionine to caffeic acid and its CoA ester to form ferulic acid and its CoA ester (Inoue et al., 1998) (Fig. 5.4). Many related genes have been identified from different plants (Table 5.3). Down regulation of COMT and CCoAOMT in transgenic plants modified the structure and composition of lignin in tobacco, poplar, alfalfa and *Arabidopsis* (Marita et al., 2003; Do et al., 2007).

cDNA sequences encoding CCoAOMT and COMT from higher plants (Table 5.3) were retrieved from the NCBI databases and partial sequences were excluded. Multiple sequence alignments were performed to generate a Neighbour-Joining tree for CCoAOMT and COMT (Fig.5.8). The phylogenetic tree for CCoAOMT and COMT was rooted to AF064696 from *Thalictrum tuberosum*. The tree was designed to determine the relationship of these genes to each other. The tree does not show separate clusters of monocotyledons and dicotyledons genes. This could be due to gene duplication occurring prior to the separation of monocotyledons and dicotyledons. The

tree shows high similarity of genes in the same plant species for example all Zea genes are in one cluster.

Table 5.3

Caffeoyl-CoA *O*-methyltransferase (CCoAOMT)(partial codons and 60 *zea mays* genes excluded because they are several copies of the same gene in the genome)

Accession number	Source	Reference
AY620245	*Ammi majus*	Unpublished
AF360317	*Arabidopsis thaliana*	Unpublished
NM_118755	*Arabidopsis thaliana*	Unpublished
NM_202375	*Arabidopsis thaliana*	Unpublished
NM_105469	*Arabidopsis thaliana*	Unpublished
NM_105468	*Arabidopsis thaliana*	Unpublished
NM_102310	*Arabidopsis thaliana*	Unpublished
AY088577	*Arabidopsis thaliana*	(Alexandrov et al., 2006)
AY088274	*Arabidopsis thaliana*	(Alexandrov et al., 2006)
AY087981	*Arabidopsis thaliana*	(Alexandrov et al., 2006)
AY087244	*Arabidopsis thaliana*	(Alexandrov et al., 2006)
AY860952	*Betula platyphylla* (Asian white birch)	Unpublished
AY651026	*Boehmeria nivea* (ramie)	Unpublished
EF015489	*Brassica rapa*	Unpublished
DQ457404	*Brassica rapa*	(Zhang et al., 2007)
AY827098	*Brassica napus*	Unpublished
AB035144	*Citrus natsudaidai*	Unpublished
EF153933	*Coffea canephora*	(Lepelley et al., 2007)
AY500159	*Corchorus capsularis*	Unpublished
AF168780	*Eucalyptus globules* (blue gum)	Unpublished
AF046122	*Eucalyptus globulus*	Unpublished
AB267815	*Ipomoea nil* (Japanese morning glory)	Unpublished
DQ517929	*Leucaena leucocephala* (white popinac)	Unpublished
DQ517930	*Leucaena leucocephala*	Unpublished
DQ431234	*Leucaena leucocephala*	Unpublished
DQ090002	*Linum usitatissimum* (flax)	Unpublished
U20736	*Medicago sativa*	(Inoue et al., 1998)
AY145521	*Mesembryanthemum crystallinum*	(Ibdah et al., 2003)
Z56282	*Nicotiana tabacum* (tobacco)	Unpublished
U62735	*Nicotiana tabacum*	(Martz et al., 1998)
Z82982	*Nicotiana tabacum*	Unpublished
AF060180	*Nicotiana tabacum*	(Maury et al., 1999)
AB110168	*Oryza sativa*	(Moriguchi et al., 2005)

NM_001070030	*Oryza sativa*	(Ohyanagi et al., 2006)
NM_001068679	*Oryza sativa*	(Ohyanagi et al., 2006)
AY644638	*Oryza sativa*	Unpublished
AM159091	*Plantago major*	Unpublished
Z33878	*Petroselinum crispum*	Unpublished
M69184	*Petroselinum crispum*	(Schmitt et al., 1991)
EF153198	*Populus deltoides*	Unpublished
U27116	*Populus tremuloides*	Unpublished
AJ224894	*Populus trichocarpa*	(Meyermans et al., 2000)
AJ224896	*Populus trichocarpa*	Unpublished
AF036095	*Pinus taeda*	(Li et al., 1999)
EU161983	*Solanum lycopersicum* (tomato)	Unpublished
AB061268	*Solanum tuberosum* (potato)	Unpublished
AF064696	*Thalictrum tuberosum*	(Frick and Kutchan, 1999)
EF178296	*Vitis vinifera* (wine grape)	Unpublished
Z54233	*Vitis vinifera*	(Busam et al., 1997)
EF535148	*Vitis vinifera*	Unpublished
AY279013	*Zea mays* (maize)	(Guillet-Claude et al., 2004)
AJ242980	*Zea mays*	Unpublished
AY279034	*Zea mays*	(Guillet-Claude et al., 2004)
AJ242981	*Zea mays*	Unpublished

Esculetin O-methyltransferase

A methyltransferase, was isolated by Ahn and co-workers from poplar leaves and then expressed in *E. coli* (Kim et al., 2006a). This gene showed high activity to the transfer of methyl groups to esculetin and generated scopoletin, isoscopoletin and scoparone (Kim et al., 2006a). The activity of catechol *O*-methyltransferase from human liver was detected by adding a methyl group to esculetin forming scopoletin (Wang et al., 2003). This step was investigated in cassava root during PPD in Chapter 4. I established that in cassava there was only a small leak of methylation from esculetin to scopoletin.

o-Hydroxylase

Although *ortho*-hydroxylation of cinnamic (*p*-coumaric, caffeic or ferulic) acids is of pivotal importance for the biosynthesis of all coumarins, it remains a missing link in phenylpropanoid biosynthesis (Bourgaud et al., 2006). This step has been discussed above in Chapter 4.

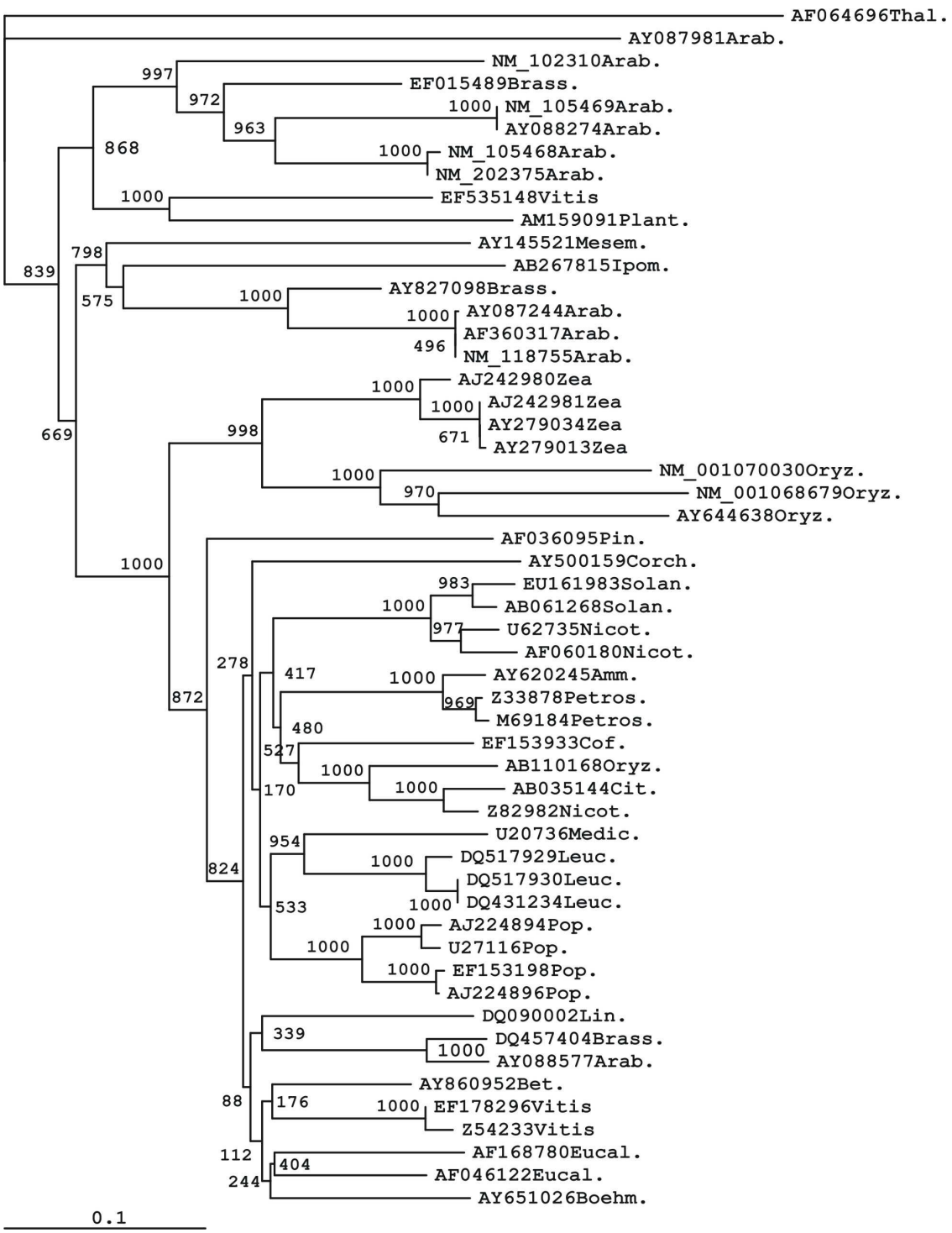


Fig. 5.8. Neighbour-Join bootstrap tree illustrating sequence similarity relationships of DNA sequences of caffeoyl-CoA *O*-methyltransferases (CCoAOMT) and caffeic acid *O*-methyltransferase (COMT) in plants.

Characterisation of relationships between plant glucosyltransferases

Plants accumulate a large range of glycosides and consequently many glycosyl transferases (GTs) have been identified. 120 GTs have been identified in *A. thaliana* (Paquette et al., 2003) and 165 in *Medicago truncatula* (Achnine et al., 2005). However, a limited number of GTs have been identified in cassava. Therefore, attempts were made to identify and isolate GTs from cassava, in particular those that might be involved in scopolin biosynthesis. This was conducted in order to characterise glucosyltransferase(s) involved in the biosynthesis of scopolin from scopoletin in cassava root during PPD.

One strategy would be to identify conserved regions amongst GTs from multiple alignments that could provide tools to isolate cassava GTs. To this end, amino acid sequences of GTs (Tables 5.4-5.8) were retrieved from the NCBI databases; these were mostly from higher plants (e.g. *M. esculenta, N. tabacum*, *A. thaliana* and *Triticum aestivum*) especially coumarin, flavonoid and anthocyanin glucosyltransferases, although a few from bacteria and yeast were also included. Flavonoid and anthocyanin GTs were chosen in addition to coumarin GTs due to the chemical structural similarity between flavonoids, anthocyanins and coumarins.

Table 5.4
Group 1 of glucosyltransferases in phylogenetic tree (Fig. 5.9)

Accession number	Source	Function	Reference	Gi number
AAR01231	*A. thaliana*	flavonoid	(Willits et al., 2004)	37703732
CAA54610	*M. esculenta*	similar to flavonoid	(Hughes and Hughes, 1994)	458547
CAB56231	*Dorotheanthus bellidiformis*	Betanidin & flavonoid	(Vogt et al., 1999)	5918023
AAS94329	*Beta vulgaris*	flavonoid	unpublished	46430995
AAB36653	*N. tabacum*	coumarin	(Horvath and Chua, 1996)	1685005
AAK28303	*N. tabacum*	phenylpropanoid	(Fraissinet-Tachet et al., 1998)	13492674

AAB36652	*N. tabacum*	coumarin	(Horvath and Chua, 1996)	1685003
BAD93688	*N. tabacum*	flavonoid	unpublished	62241063
BAD89034	*Solanum tuberosum*	steroidal saponin	(Kohara et al., 2005)	58430480
BAD89042	*Solanum aculeatissimum*	steroidal saponin	(Kohara et al., 2005)	58430496
BAD89043	*S. aculeatissimum*	steroidal saponin	(Kohara et al., 2005)	58430498
BAD89041	*Solanum khasianum*	steroidal saponin	(Kohara et al., 2005)	58430494

Table 5.5
Group 2 of glucosyltransferases in phylogenetic tree (Fig. 5.9)

Accession number	**Source**	**Function**	**Reference**	**Gi number**
AAM61455	*A. thaliana*	flavonoid	unpublished	21537114
CAB62443	*A. thaliana*	-	unpublished	6561977
NP_179446	*A. thaliana*	-	unpublished	30680413
CAA54612	*M. esculenta*	similar to flavonoid	(Hughes and Hughes, 1994)	453249
CAB16822	*A. thaliana*	-	unpublished	4006892
CAB80343	*A. thaliana*	-	unpublished	7270626
AAM61455	*A. thaliana*	flavonoid	unpublished	21537114
Q9M156	*A. thaliana*	similar to hydroquinone	(Mayer et al., 1999)	28380085
CAB809166	*A. thaliana*	flavonoid	unpublished	7267604
AAK25972	*A. thaliana*	flavonoid	unpublished	13430700
B85014	*A. thaliana*	Similar to flavonoid	(Mayer et al., 1999)	25286798
AAK64133	*A. thaliana*	Similar to flavonoid	unpublished	14532902

Table 5.6
Group 3 of glucosyltransferases in phylogenetic tree (Fig. 5.9)

Accession number	Source	Function	Reference	Gi number
NP_563784	*A. thaliana*	-	unpublished	30680040
AGG48783	*A. thaliana*	-	unpublished	12083248
AAM65993	*A. thaliana*	-	unpublished	21594310
AAS94330	*Beta vulgaris*	flavonoid	unpublished	46430997
AAU09444	*Fragaria x ananassa*	flavonoid	unpublished	51705413
CAA54609	*M. esculenta*	similar to flavonoid	(Hughes and Hughes, 1994)	453246
CAA54611	*M. esculenta*	similar to flavonoid -	(Hughes and Hughes, 1994)	453255
CAA54613	*M. esculenta*	similar to flavonoid -	(Hughes and Hughes, 1994)	453251
BAB88934	*Nicotiana tabacum*	-	(Taguchi et al., 2003a)	20146091
BAB60721	*N. tabacum*	-	(Taguchi et al., 2001)	14349253
BAB60720	*N. tabacum*	-	(Taguchi et al., 2001)	14349251

Table 5.7
Group 4 of glucosyltransferases in phylogenetic tree (Fig. 5.9)

Accession number	Source	Function	Reference	Gi number
CAC01718	*A. thaliana*	flavonoid	unpublished	9755706
CAC01717	*A. thaliana*	flavonoid	unpublished	9755705
CAC01716	*A. thaliana*	flavonoid	unpublished	9755704
AAS89832	*Fragaria x ananassa*	flavonoid	unpublished	46370000
CAA33729	*Hordeum vulgare*	flavonoid	(Wise et al., 1990)	295807
BAA90787	*Ipomoea batatas*	flavonoid	unpublished	6983839
CAA54614	*M. esculenta*	similar to flavonoid	(Hughes and Hughes, 1994)	453253
BAA19659	*Perilla frutescens*	flavonoid	unpublished	1944201
BAB41026	*Vitis vinifera*	flavonoid	(Kobayashi et al., 2001)	13620873
BAB41022	*V. vinifera*	flavonoid	(Kobayashi et al., 2001)	13620865

Table 5.8
Other glucosyltransferases in phylogenetic tree (Fig. 5.9)

Accession number	Source	Function	Refrence	Gi number
AAO64763	*A. thaliana*	-	unpublished	29028768
AAD38266	*A. thaliana*	flavonoid	unpublished	5042427
AAD38265	*A. thaliana*	flavonoid	unpublished	5042426
AAM64627	*A. thaliana*	-	unpublished	21592678
AAM64890	*A. thaliana*	-	unpublished	21592940
AAO11554	*A. thaliana*	-	unpublished	27363270
AAQ06264	*S. bicolor*	flavonoid	unpublished	33321017
AAQ06263	*S. bicolor*	flavonoid	unpublished	33321016
BAC91654	*Gloeobacter violaceus*	-	(Nakamura et al., 2003)	35214286
BAC78438	*Glycyrrhiza echinata*	isoflavonoid	(Nagashima et al., 2004)	32188025
BAB41023	*V. vinifera*	flavonoid	unpublished	13620867
BAC91659	*G. violaceus*	-	(Nakamura et al., 2003)	35214291
BAA83484	*Scutellaria baicalensis*	flavonoid	(Hirotani et al., 2000)	5763524
AAF17077	*Sorghum bicolor*	Similar to cyanogenic glycosides	(Jones et al., 1999)	6561805
CAD27851	*T. aestivum*	-	unpublished	57282068
CAD27849	*T. aestivum*	-	unpublished	57282064
CAD27847	*T. aestivum*	-	unpublished	23392915
CAD27850	*T. aestivum*	-	unpublished	57282066
CAD27857	*T. aestivum*	-	unpublished	23392925
CAD27859	*T. aestivum*	-	unpublished	23392929
CAD27858	*T. aestivum*	-	unpublished	23392927
CAD27860	*T. aestivum*	-	unpublished	23392937
CAD27852	*T. aestivum*	-	unpublished	57282070
CAD27851	*T. aestivum*	-	unpublished	57282068
BAB88935	*N. tabacum*	flavonoid and coumarin	(Taguchi et al., 2003b)	20146093
BAA36423	*Verbena x hybrida*	anthocyanin	(Yamazaki et al., 1999)	4115563
BAA36422	*Perilla frutescens*	anthocyanin	(Yamazaki et al., 1999)	4115561
BAA36421	*P. frutescens*	anthocyanin	(Yamazaki et al., 1999)	4115559
AAF61647	*N. tabacum*	salicylic acid	(Lee and Raskin, 1999)	7385017

P50076	*Saccharomyces cerevisiae*	Asparagine-linked glycosylation protein	(Nikawa and Hosaka, 1995)	1706435
BAB41025	*V. vinifera*	flavonoid	(Kobayashi et al., 2001)	13620871
AAM62706	*A. thaliana*	-	unpublished	21553613
AAB62270	*Solanum berthaultii*	-	unpublished	2232354
NP_486328	*Nostoc*	-	(Kaneko et al., 2001)	17229780
BAB41024	*V. vinifera*	flavonoid	(Kobayashi et al., 2001)	13620869
AAT42163	*S. bicolor*	zeatin	(Swigonova et al., 2004)	48374965
AAU90273	*Oryza sativa*	hydroquinone	unpublished	53749415
BAA12737	*Gentiana triflora*	flavonoid	(Tanaka et al., 1996)	1620013

Multiple sequence alignments were performed (using CLUSTAL X) and used to generate a Neighbour-Joining tree (Fig. 5.9) using 1000 bootstrap trials and rooted to glucosyltransferase from *Xanthomonas campestris* pv. *campestris* (YP_245010). The tree was designed to determine the relationship of GTs to each other and their relation to cassava GTs. The neighbour Joining tree (Fig. 5.9) shows distinct clusters of related GTs including cassava GTs, though these latter are distributed throughout the tree rather than clustered in one group. GTs with the same function are clustered in groups, in particular if they are from the same genus e.g. flavonoid GTs from *A. thaliana* in group 2, but other flavonoid GT from a different genus could be related e.g. CAA54612 which is a flavonoid GT from *M. esculenta.* Five related groups were chosen for further analysis in addition to steroidal saponin GTs from different species of *Solanum*. Group 1 includes *N. tabacum* coumarin GTs and flavonoid GTs from *Beta vulgaris*, *A. thaliana*, *Dorotheanthus bellidiformis* and *M. esculenta* confirming the relationship of flavonoid and coumarin GTs. Group 2 includes mainly flavonoid GTs from *A. thaliana* and one GT from *M. esculenta*. Group 3 includes three related GTs from *M. esculenta,* flavonoid GTs from *A. thaliana* and *Fragaria x ananassa* and other GTs of unknown function from *A. thaliana* and *N. tabacum*. Group 4 includes flavonoid GTs from different genera, *A. thaliana*, *Fragaria x ananassa*, *Hordeum vulgare*, *Ipomoea batatas*, *M. esculenta*, *Perilla frutescens* and *Vitis vinifera*.

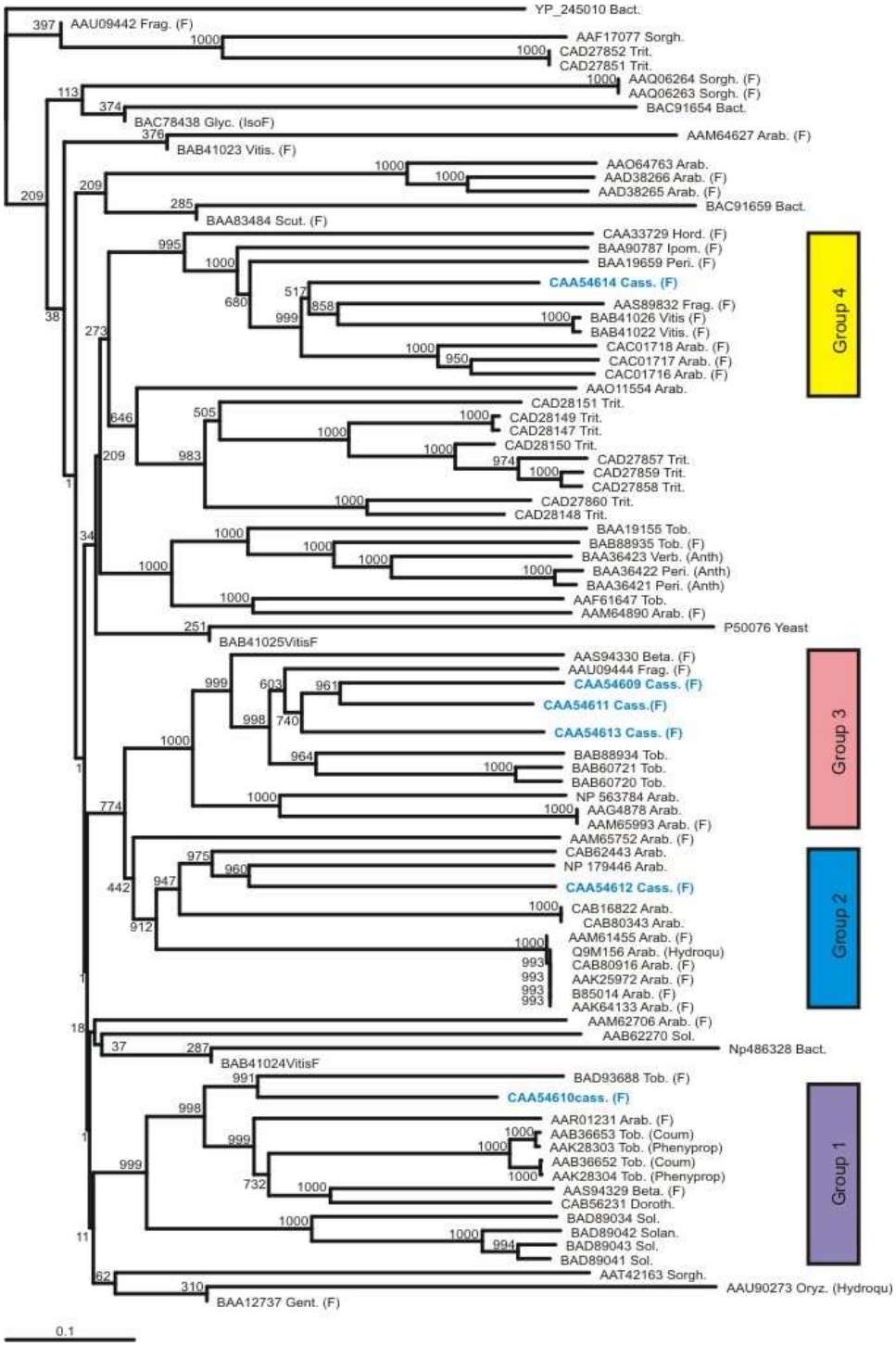
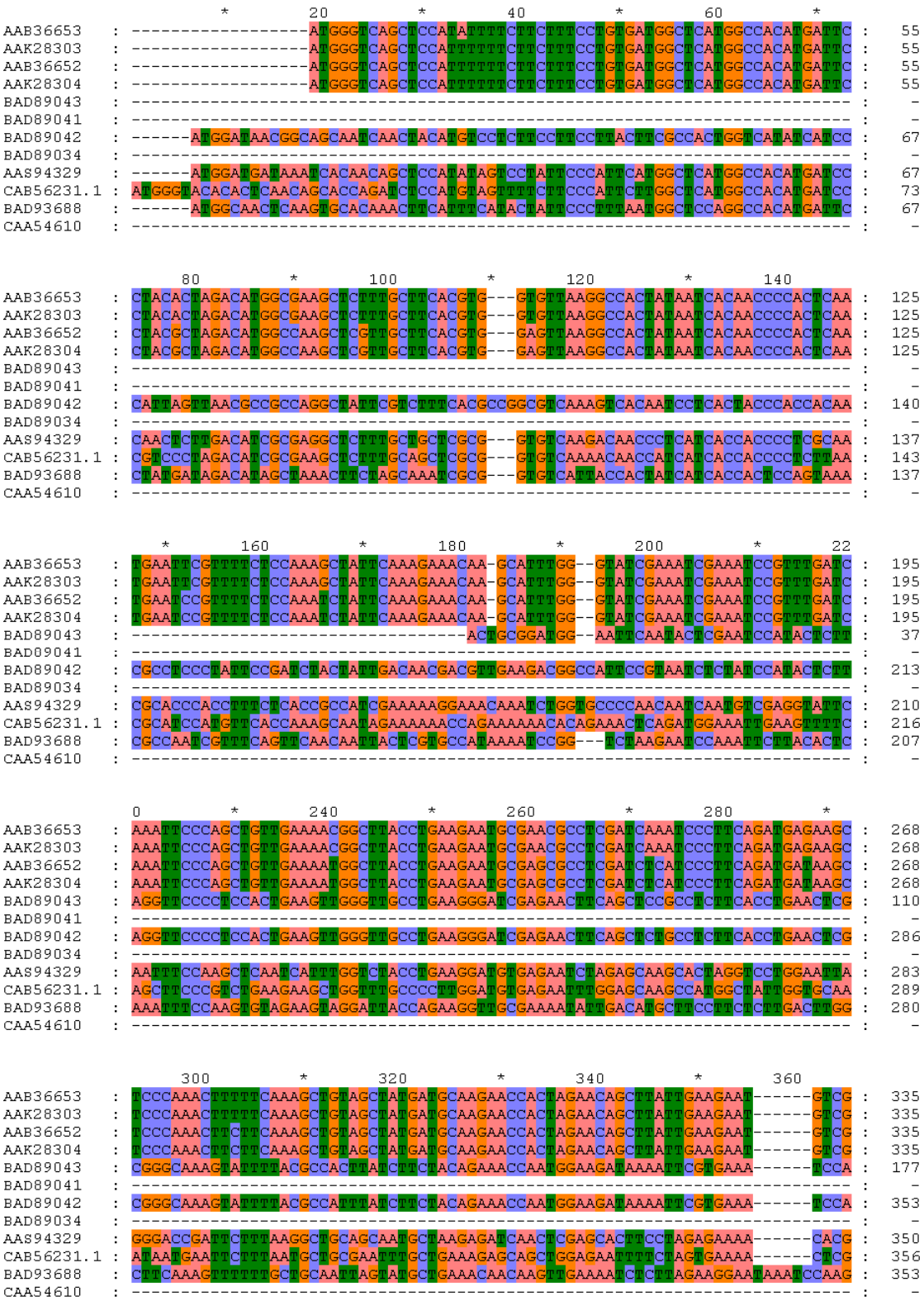
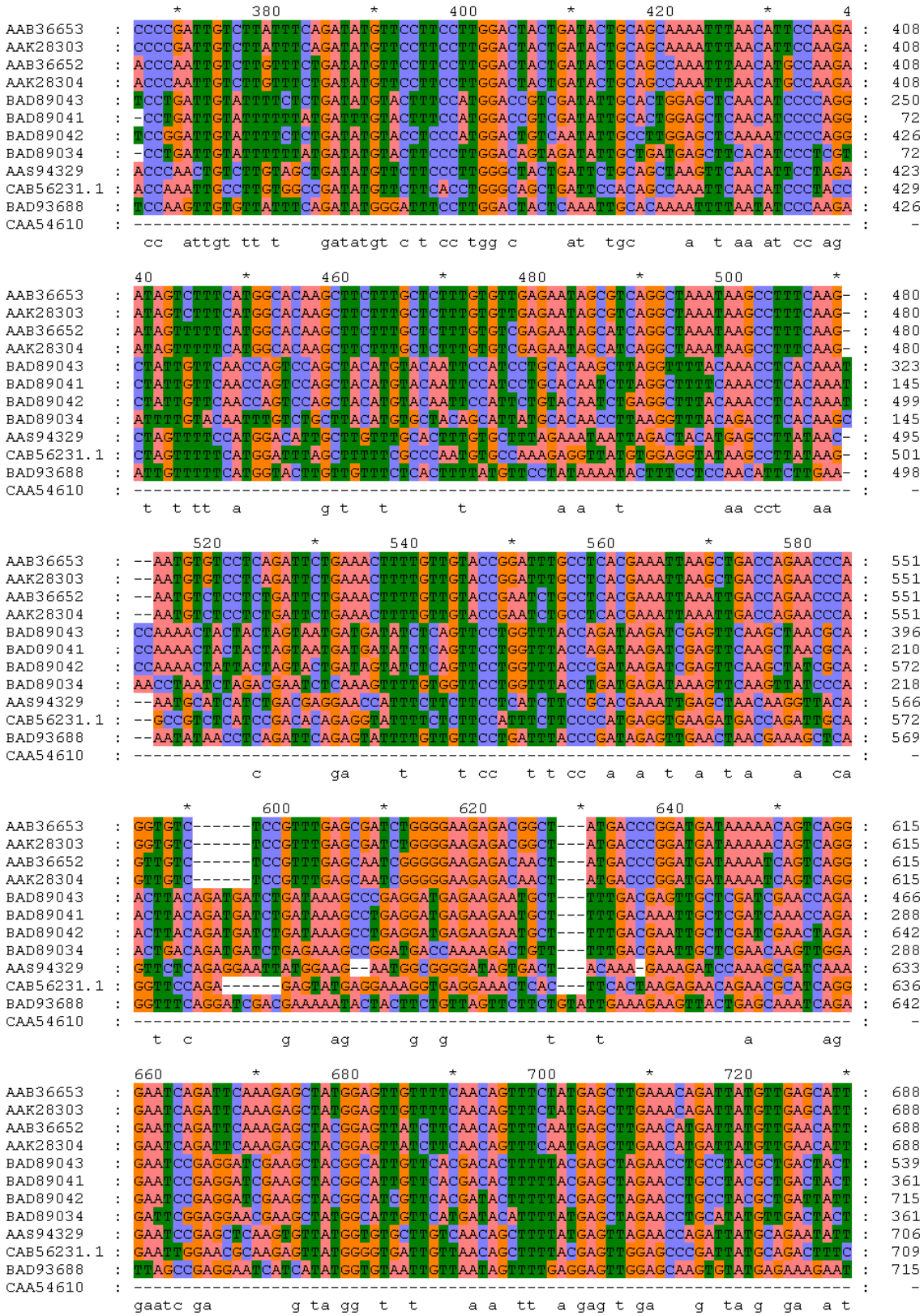


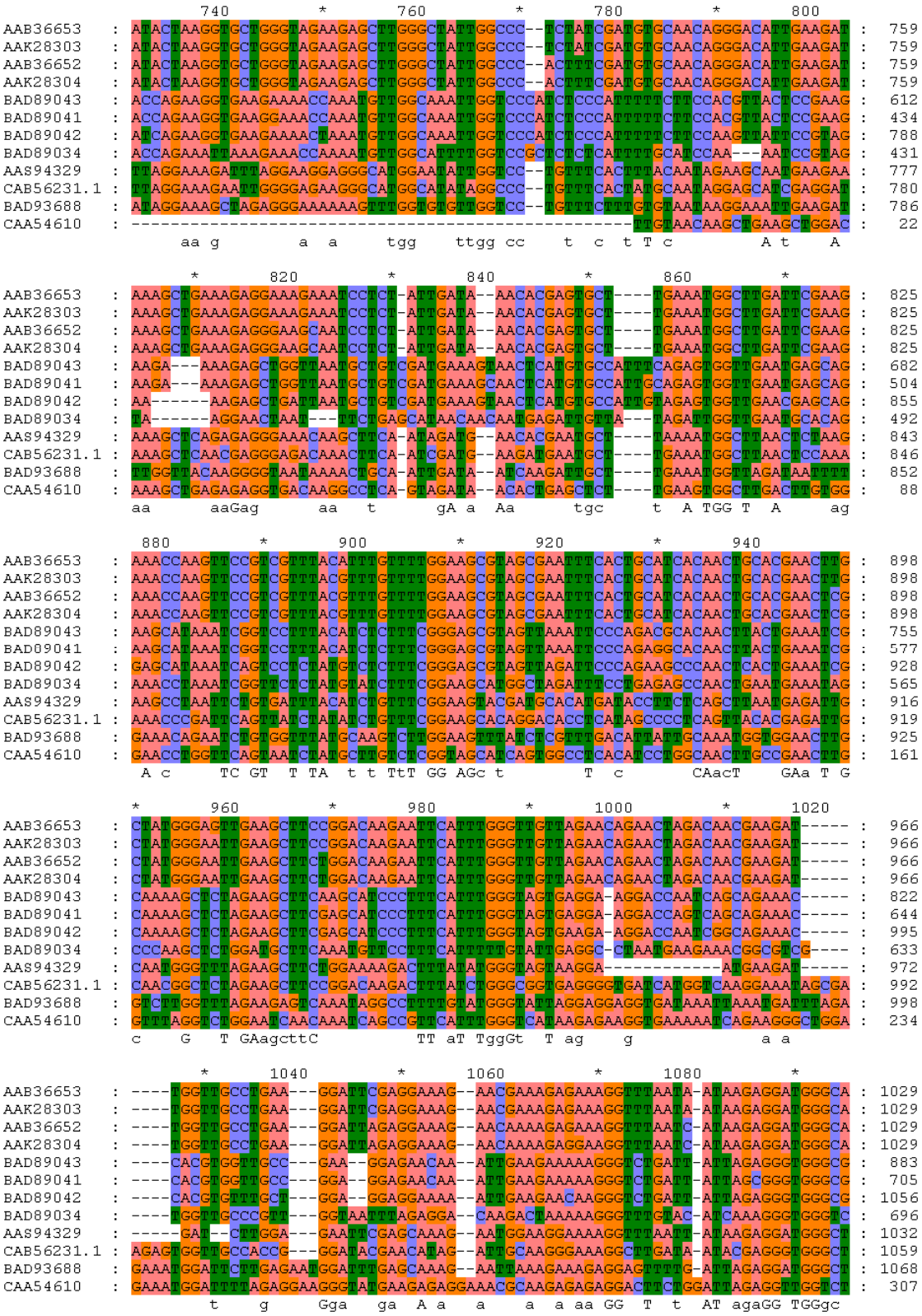
Fig. 5.9. Neighbour-Joining bootstrap tree (designed using CLUSTAL X) of different glucosyltransferases from various plants. Abbreviations, Anthocyanin GT (Anth), Coumarin (Coum), Flavonoid (F), Hydroquinone (Hydroqu) and Phenylpropanoid (Phenylprop). Groups 1–4 are discussed in the text. Cassava GTs are coloured blue.

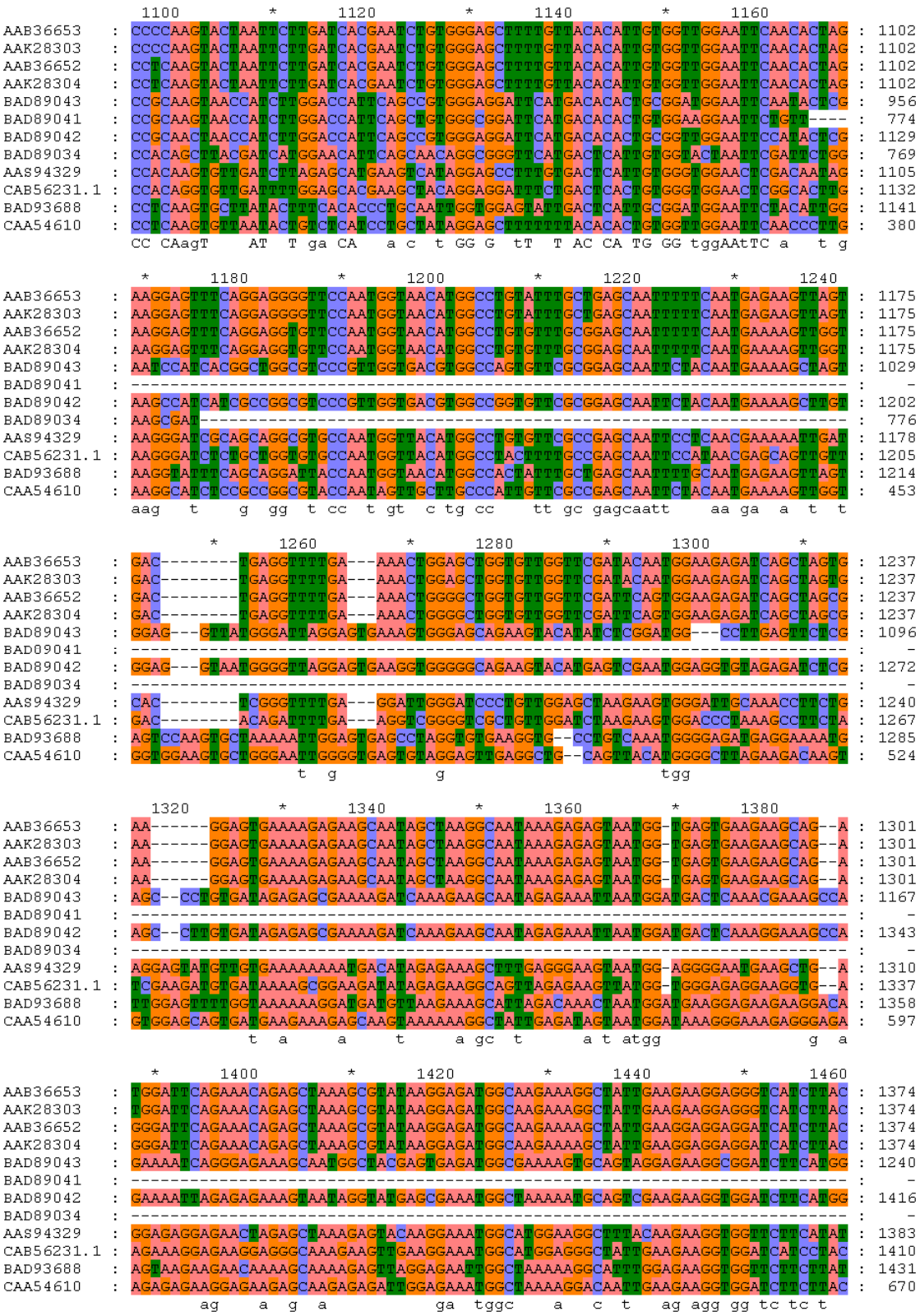
Design of PCR primers to isolate representatives of glucosyltransferases from cassava

As a result of studying this tree, four pairs of degenerate primers were designed specific to four chosen related groups instead of designing a pair of universal primers for all glucosyltransferases. Multiple alignments of nucleotide sequences of each of the four groups of glucosyltransferases (Table 5.4-5.7) were carried out (Fig. 5.10, 5.12, 5.14 and 5.16). Some glucosyltransferase sequences were omitted from the alignment in order to increase the conserved regions in the alignment. Depending on the conserved regions, both (glucosyl transferase) forward and reverse degenerate primers GTF1-4 GTR1-4 were designed for the purpose of amplifying and isolating representative members of each group from cassava PPD-related cDNA libraries (Fig. 5.11, 5.13, 5.15 and 5.17) (Table 5.9).









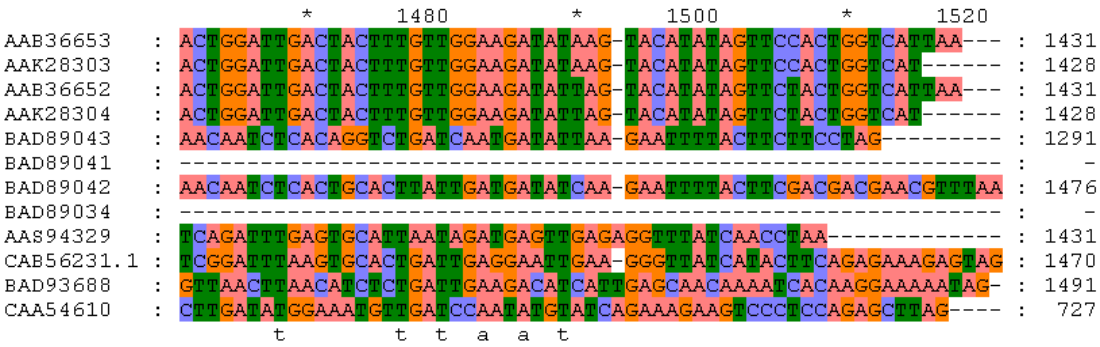


Fig. 5.10. Nucleotide alignments of group 1 glucosyltransferases.

In group 1 alignment (Fig. 5.10), the steroidal saponin GTs BAD89041 and BAD89034 from *Solanum* as well as CAA54610 from *M. esculenta* have shorter sequences than other GTs in the group, so these three sequences were omitted from the alignment then alignment was performed. This increased the conserved region in the alignment to design the forward and reverse degenerate primers GTF1 and GTR1 (Fig. 5.11).

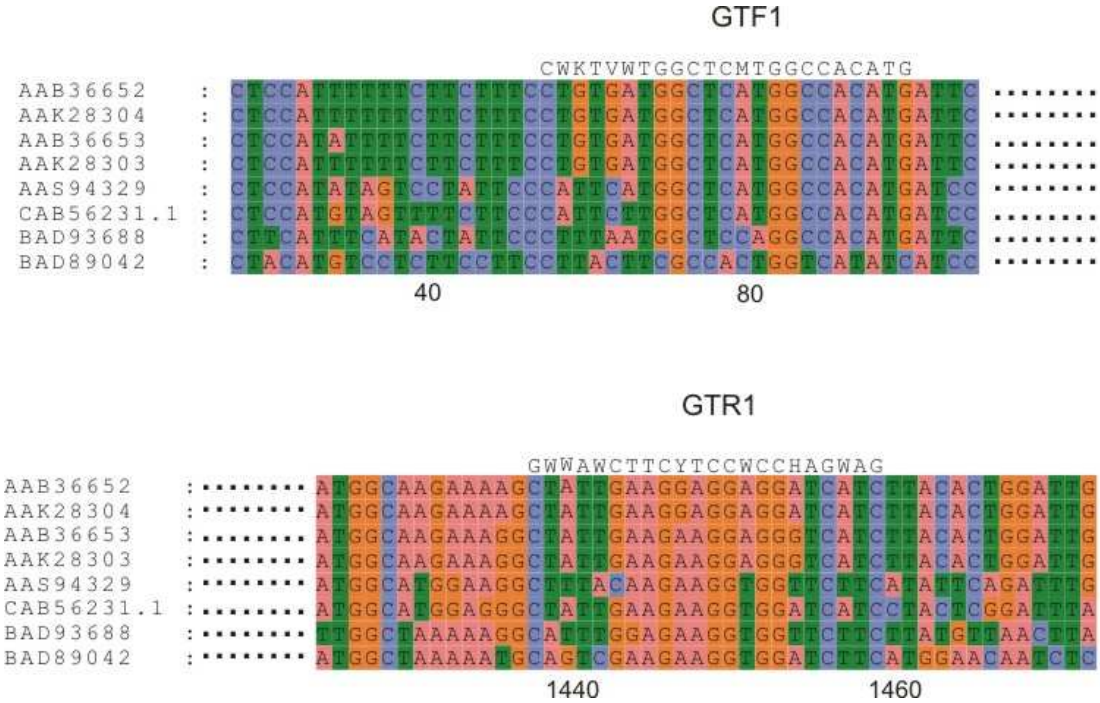
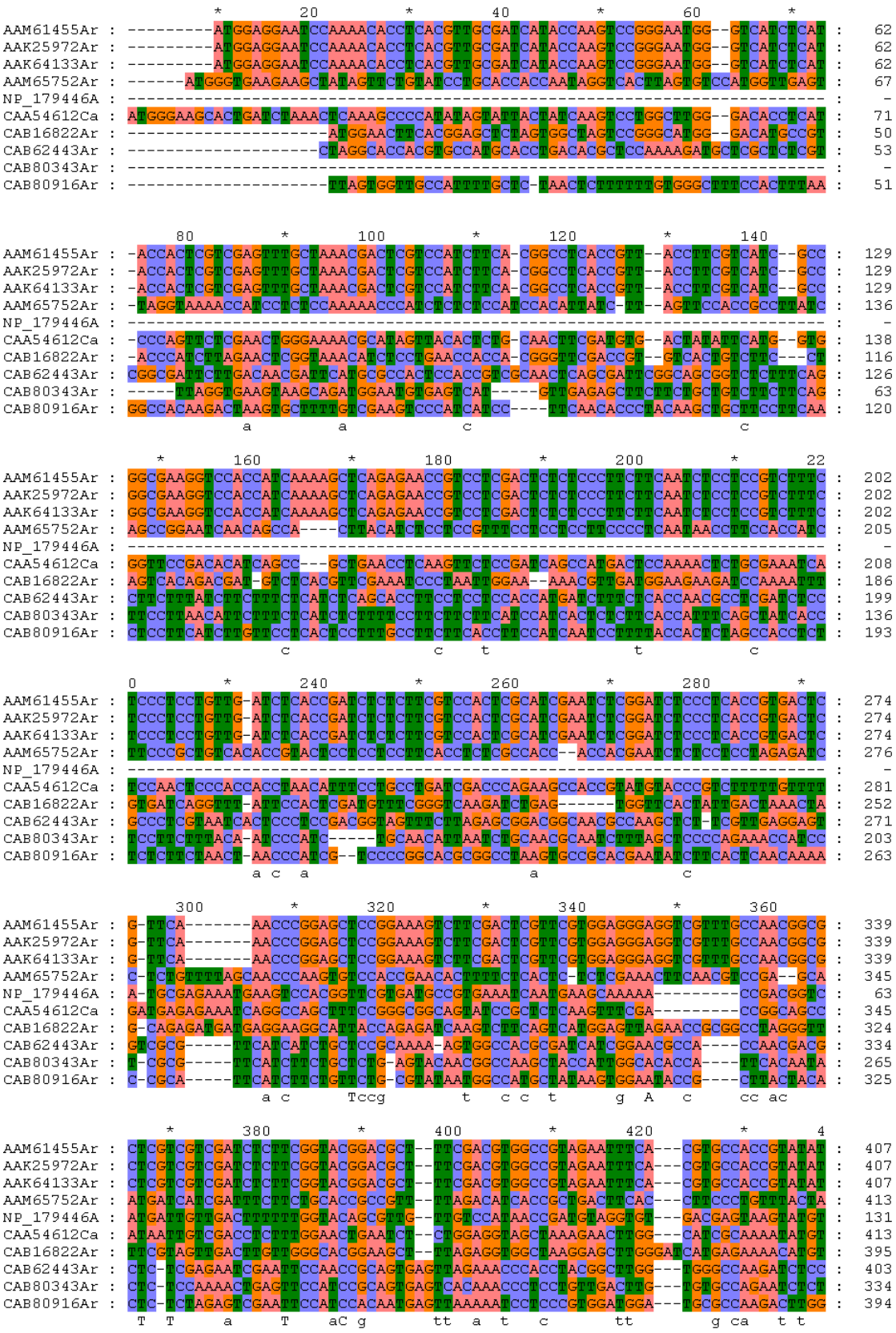
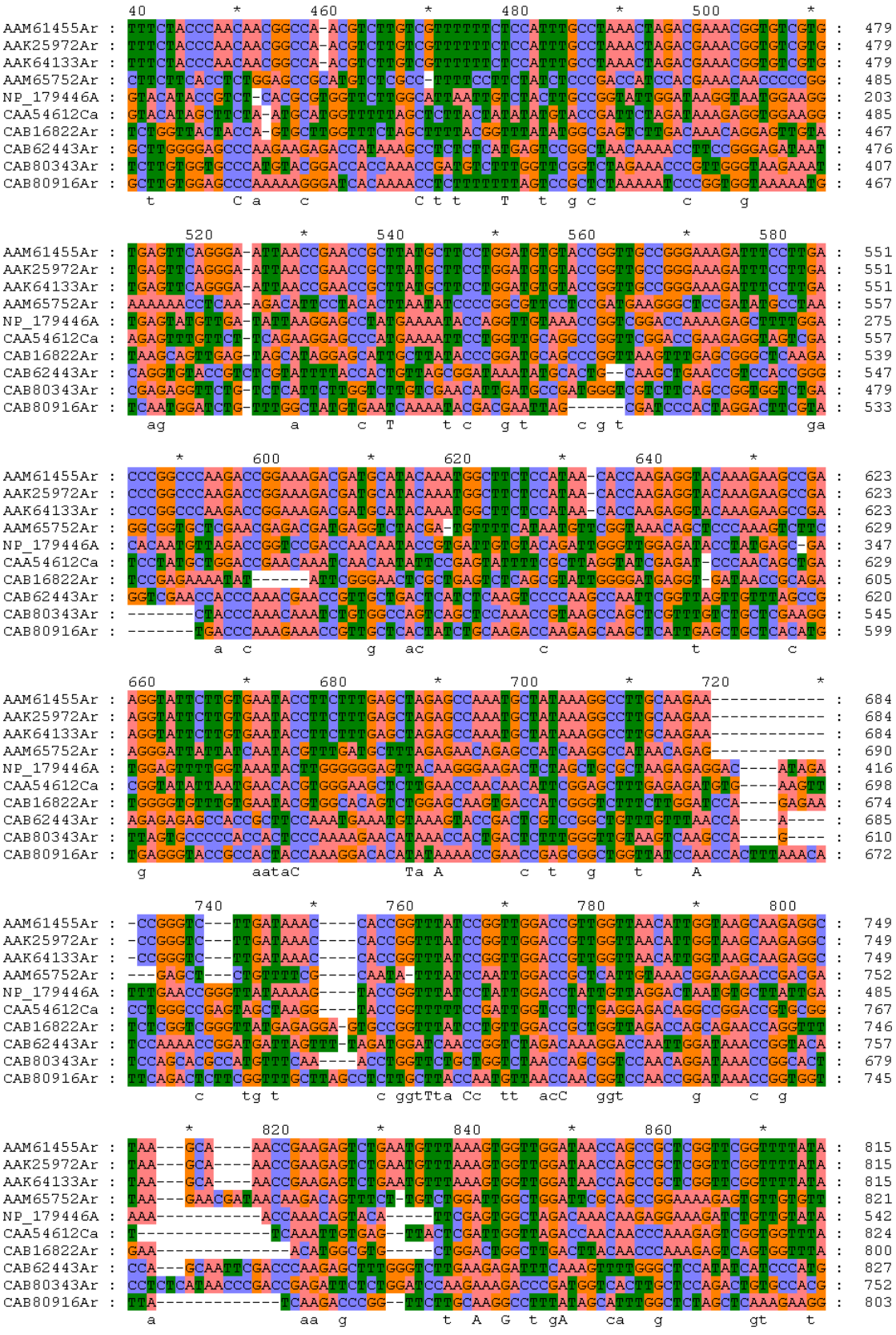
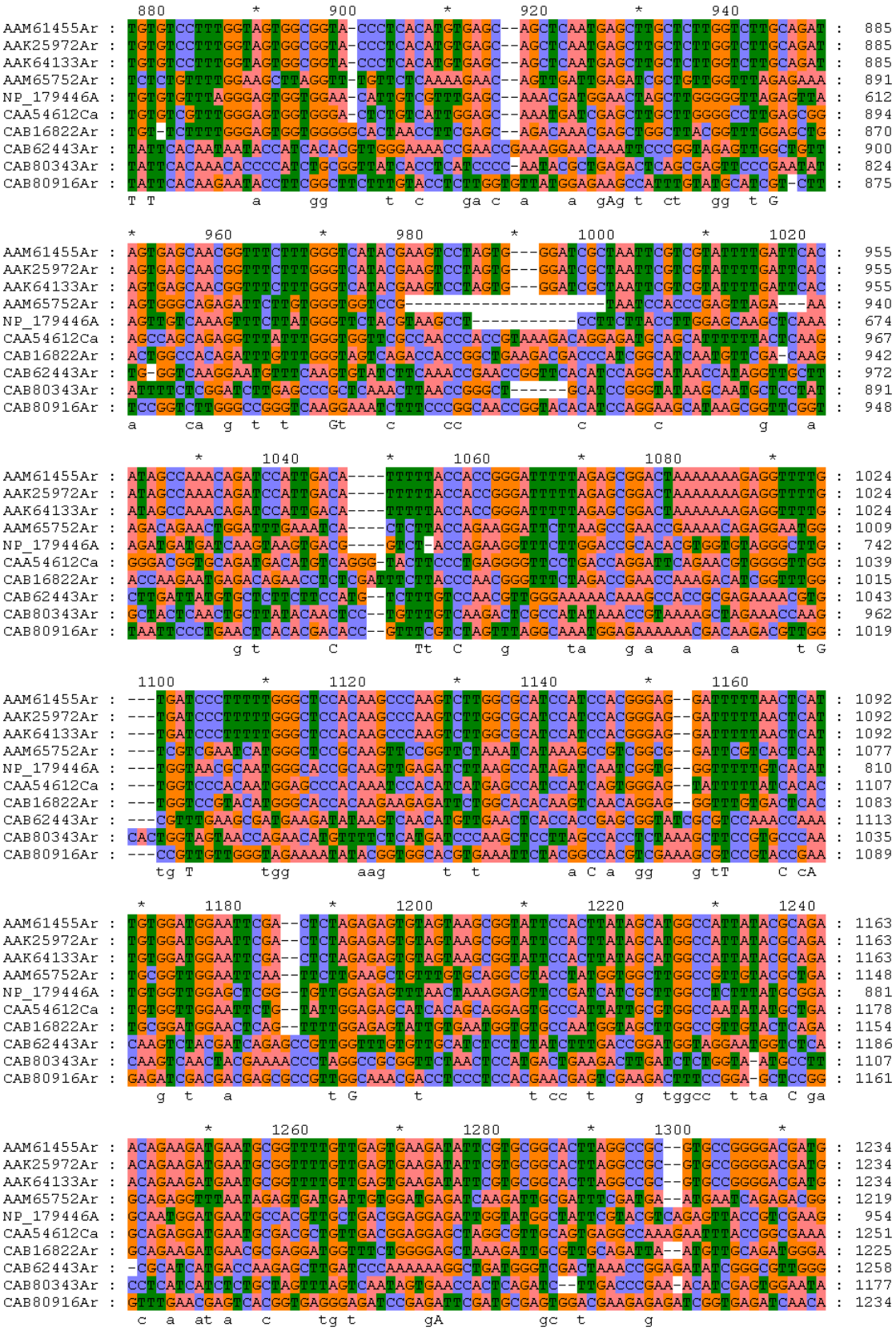


Fig. 5.11. Nucleotide alignment of isolates of group 1 GT. Amplification primers (GTF1/R1) are indicated. 5', 3' and core sequences have been omitted for illustration purposes. Abbreviations: Adenine (A), Cytosine (C), Guanine (G), Thymine (T), B (G+T+C), D (G+A+T), H (A+T+C), K (G+T), M (A+C), N (A+C+G+T), V (G+A+C), W (A+T), R (A+G), Y (C+T)






880 * 900 * 920 * 940
AAM61455Ar : TGTGTCCTTTGGTAGTGGCGGTA-CCCTCACATGTGAGC--AGCTCAATGAGCTTGCTCTTGGTCTTGCAGAT : 885
AAK25972Ar : TGTGTCCTTTGGTAGTGGCGGTA-CCCTCACATGTGAGC--AGCTCAATGAGCTTGCTCTTGGTCTTGCAGAT : 885
AAK64133Ar : TGTGTCCTTTGGTAGTGGCGGTA-CCCTCACATGTGAGC--AGCTCAATGAGCTTGCTCTTGGTCTTGCAGAT : 885
AAM65752Ar : TCTCTGTTTTGGAAGCTTAGGTT-TGTTCTCAAAAGAAC--AGTTGATTGAGATCGCTGTTGGTTTAGAGAAA : 891
NP_179446A : TGTGTGTTTAGGGAGTGGTGGAA-CATTGTCGTTTGAGC--AAACGATGGAACTAGCTTGGGGGTTAGAGTTA : 612
CAA54612Ca : TGTGTCGTTTGGGAGTGGTGGGA-CTCTGTCATTGGAGC--AAATGATCGAGCTTGCTTGGGGCCTTGAGCGG : 894
CAB16822Ar : TGT-TCTTTTGGGAGTGGTGGGGGCACTAACCTTCGAGC--AGACAAACGAGCTGGCTTACGGTTTGGAGCTG : 870
CAB62443Ar : TATTCACAATAATACCATCACACGTTGGGAAAACCGAACCGAAAGGAACAAATTCCCGGTAGAGTTGGCTGTT : 900
CAB80343Ar : TATTCACAAACACCCCATCTGCGGTTATCACCTCATCCCC-AATACGCTGAGACTCAGCGAGTTCCCGAATAT : 824
CAB80916Ar : TATTCACAAGAATACCTTCGGCTTCTTTGTACCTCTTGGTGTTATGGAGAAGCCATTTGTATGCATCGT-CTT : 875
T T a gg t c ga c a a gAg t ct gg t G
* 960 * 980 * 1000 * 1020
AAM61455Ar : AGTGAGCAACGGTTTCTTTGGGTCATACGAAGTCCTAGTG---GGATCGCTAATTCGTCGTATTTTGATTCAC : 955
AAK25972Ar : AGTGAGCAACGGTTTCTTTGGGTCATACGAAGTCCTAGTG---GGATCGCTAATTCGTCGTATTTTGATTCAC : 955
AAK64133Ar : AGTGAGCAACGGTTTCTTTGGGTCATACGAAGTCCTAGTG---GGATCGCTAATTCGTCGTATTTTGATTCAC : 955
AAM65752Ar : AGTGGGCAGAGATTCTTGTGGGTGGTCCG----------TAATCCACCCGAGTTAGA---AA : 940
NP_179446A : AGTTGTCAAAGTTTCTTATGGGTTCTACGTAAGCCT----------CCTTCTTACCTTGGAGCAAGCTCAAA : 674
CAA54612Ca : AGCCAGCAGAGGTTTATTTGGGTGGTTCGCCAACCCACCGTAAAGACAGGAGATGCAGCATTTTTTACTCAAG : 967
CAB16822Ar : ACTGGCCACAGATTTGTTTGGGTAGTCAGACCACCGGCTGAAGACGACCCATCGGCATCAATGTTCGA-CAAG : 942
CAB62443Ar : TG-GGTCAAGGAATGTTTCAAGTGTATCTTCAAACCGAACCGGTTCACATCCAGGCATAACCATAGGTTGCTT : 972
CAB80343Ar : ATTTTCTCGGATCTTGAGCCCGCTCAAACTTAACCGGGCT------GCATCCGGGTATAAGCAATGCTCCTAT : 891
CAB80916Ar : TCCGGTCTTGGGCCGGGTCAAGGAAATCTTTCCCGGCAACCGGTACACATCCAGGAAGCATAAGCGGTTCGGT : 948
a ca g t t Gt c cc c c g a
* 1040 * 1060 * 1080 *
AAM61455Ar : ATAGCCAAACAGATCCATTGACA----TTTTTACCACCGGGATTTTTAGAGCGGACTAAAAAAAGAGGTTTTG : 1024
AAK25972Ar : ATAGCCAAACAGATCCATTGACA----TTTTTACCACCGGGATTTTTAGAGCGGACTAAAAAAAGAGGTTTTG : 1024
AAK64133Ar : ATAGCCAAACAGATCCATTGACA----TTTTTACCACCGGGATTTTTAGAGCGGACTAAAAAAAGAGGTTTTG : 1024
AAM65752Ar : AGACAGAACTGGATTTGAAATCA----CTCTTACCAGAAGGATTCTTAAGCCGAACCGAAAACAGAGGAATGG : 1009
NP_179446A : AGATGATGATCAAGTAAGTGACG----GTCT-ACCAGAAGGTTTCTTGGACCGCACACGTGGTGTAGGGCTTG : 742
CAA54612Ca : GGGACGGTGCAGATGACATGTCAGGG-TACTTCCCTGAGGGGTTCCTGACCAGGATTCAGAACGTGGGGTTGG : 1039
CAB16822Ar : ACCAAGAATGAGACAGAACCTCTCGATTTCTTACCCAACGGGTTTCTAGACCGAACCAAAGACATCGGTTTGG : 1015
CAB62443Ar : CTTGATTATGTGTCTTCTTCCATG--TCTTTGTCCAACGTTGGGAAAAACAAAGCCACCGCGAGAAAACGTG : 1043
CAB80343Ar : GCTACTCAACTGCTTATACAACTCC--TGTTTGTCAAGACTCGCCATATAAACCGTAAAAGCTAGAAACCAAG : 962
CAB80916Ar : TAATTCCCTGAACTCACACGACACC--GTTTCGTCTAGTTTAGGCAAATGGAGAAAAAACGACAAGACGTTGG : 1019
g t C Tt C g ta g a a a t G
1100 * 1120 * 1140 * 1160
AAM61455Ar : ---TGATCCCTTTTTGGGCTCCACAAGCCCAAGTCTTGGCGCATCCATCCACGGGAG--GATTTTTAACTCAT : 1092
AAK25972Ar : ---TGATCCCTTTTTGGGCTCCACAAGCCCAAGTCTTGGCGCATCCATCCACGGGAG--GATTTTTAACTCAT : 1092
AAK64133Ar : ---TGATCCCTTTTTGGGCTCCACAAGCCCAAGTCTTGGCGCATCCATCCACGGGAG--GATTTTTAACTCAT : 1092
AAM65752Ar : ---TCGTCGAAATCATGGGCTCCGCAAGTTCCGGTTCTAAATCATAAAGCCGTCGGCG--GATTCGTCACTCAT : 1077
NP_179446A : ---TGGTAACGCAATGGGCACCGCAAGTTGAGATCTTAAGCCATAGATCAATCGGTG--GGTTTTTGTCACAT : 810
CAA54612Ca : ---TGGTCCCACAATGGAGCCCACAAATCCACATCATGAGCCATCCATCAGTGGGAG--TATTTTTATCACAC : 1107
CAB16822Ar : ---TGGTCCGTACATGGGCACCACAAGAAGAGATTCTGGCACACAAGTCAACAGGAG--GGTTTGTGACTCAC : 1083
CAB62443Ar : ---CGTTTGAAGCGATGAAGATATAAGTCAACATGTTGAACTCACCACCGAGCGGTATCGCGTCCAAACCAAA : 1113
CAB80343Ar : CACTGGTAGTAACCAGAACATGTTTTCTCATGATCCCAAGCTCCTTAGCCACCTCTAAAGCTTCCGTGCCCAA : 1035
CAB80916Ar : ---CCGTTGTTGGGTAGAAAATATACGGTGGCACGTGAAATTCTACGGCCACGTCGAAAGCGTCCGTACCGAA : 1089
tg T tgg aag t t a C a gg g tT C cA
* 1180 * 1200 * 1220 * 1240
AAM61455Ar : TGTGGATGGAATTCGA--CTCTAGAGAGTGTAGTAAGCGGTATTCCACTTATAGCATGGCCATTATACGCAGA : 1163
AAK25972Ar : TGTGGATGGAATTCGA--CTCTAGAGAGTGTAGTAAGCGGTATTCCACTTATAGCATGGCCATTATACGCAGA : 1163
AAK64133Ar : TGTGGATGGAATTCGA--CTCTAGAGAGTGTAGTAAGCGGTATTCCACTTATAGCATGGCCATTATACGCAGA : 1163
AAM65752Ar : TGCGGTTGGAAATTCAA--TTCTTGAAGCTGTTTGTGCAGGCGTACCTATGGTGGCTTGGCCGTTGTACGCTGA : 1148
NP_179446A : TGTGGTTGGAGCTCGG--TGTTGGAGAGTTTAACTAAAGGAGTTCCGATCATCGCTTGGCCTCTTTATGCGGA : 881
CAA54612Ca : TGTGGTTGGAATTCTG--TATTGGAGAGCATCACAGCAGGAGTGCCCATTATTGCGTGGCCAATATATGCTGA : 1178
CAB16822Ar : TGCGGATGGAACTCAG--TTTTGGAGAGTATTGTGAATGGTGTGCCAATGGTAGCTTGGCCGTTGTACTCAGA : 1154
CAB62443Ar : CAAGTCTACGATCAGAGCCGTTGGTTTGTGTTGCATCTCCTCTATCTTTGACCGGATGGTAGGAATGGTCTCA : 1186
CAB80343Ar : CAAGTCAACTACGAAAACCCTAGGCCGCGGTTCTAACTCCATGACTGAAGACTTGATCTCTGGTA-ATGCCTT : 1107
CAB80916Ar : GAGATCGACGACGAGCGCCGTTGGCAAACGACCTCCCTCCACGAACGAGTCGAAGACTTTCCGGA-GCTCCGG : 1161
g t a t G t t cc t g tggcc t ta C ga
* 1260 * 1280 * 1300 *
AAM61455Ar : ACAGAAGATGAATGCGGTTTTGTTGAGTGAAGATATTCGTGCGGCACTTAGGCCGC--GTGCCGGGGACGATG : 1234
AAK25972Ar : ACAGAAGATGAATGCGGTTTTGTTGAGTGAAGATATTCGTGCGGCACTTAGGCCGC--GTGCCGGGGACGATG : 1234
AAK64133Ar : ACAGAAGATGAATGCGGTTTTGTTGAGTGAAGATATTCGTGCGGCACTTAGGCCGC--GTGCCGGGGACGATG : 1234
AAM65752Ar : GCAGAGGTTTAATAGAGTGATGATTGTGGATGAGATCAAGATTGCGATTTCGATGA--ATGAATCAGAGACGG : 1219
NP_179446A : GCAATGGATGAATGCCACGTTGCTGACGGAGGAGATTGGTATGGCTATTCGTACGTCAGAGTTACCGTCGAAG : 954
CAA54612Ca : GCAGAGGATGAATGCGACGCTGTTGACGGAGGAGCTAGGCGTTGCACTGAGGCCAAAGAATTTACCGGCGAAA : 1251
CAB16822Ar : GCAGAAGATGAACGCGAGGATGGTTTCTGGGGAGCTAAAGATTGCGTTGCAGATTA--ATGTTGCAGATGGGA : 1225
CAB62443Ar : -CGCATCATGACCAAGAGCTTGATCCCAAAAAAGGCTGATGGGTCGACTAAACCGGAGATATCGGGCGTTGGG : 1258
CAB80343Ar : CCTCATCATCTCTGCTAGTTTAGTCAATAGTGAACCACTCAGATC--TTGACCCGAA-ACATCGAGTGGAATA : 1177
CAB80916Ar : GTTTGAACGAGTCACGGTGAGGGAGATCCGAGATTCGATGCGAGTGGACGAAGAGAGATCGGTGAGATCAACA : 1234
c a at a c tg t gA gc t g

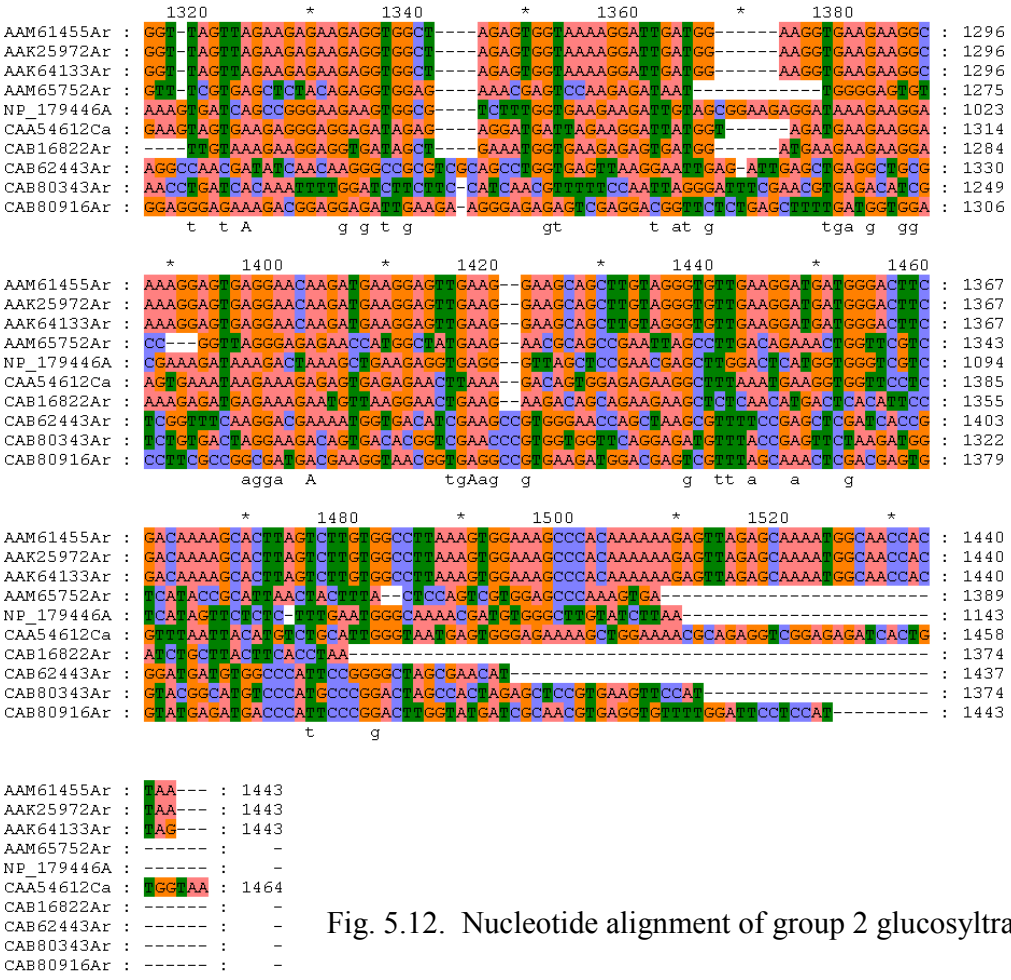


Fig. 5.12. Nucleotide alignment of group 2 glucosyltransferases.

Group 2 nucleotide sequences alignment (Fig. 5.12) did not show enough conserved regions. Only three related sequences from *A. thaliana* in addition to CAA54612 which is a flavonoid GT from *M. esculenta* were used. Degenerate forward and reverse primersGTF2 and GTF3 were designed from chosen conserved regions along the alignment (Fig. 5.13)

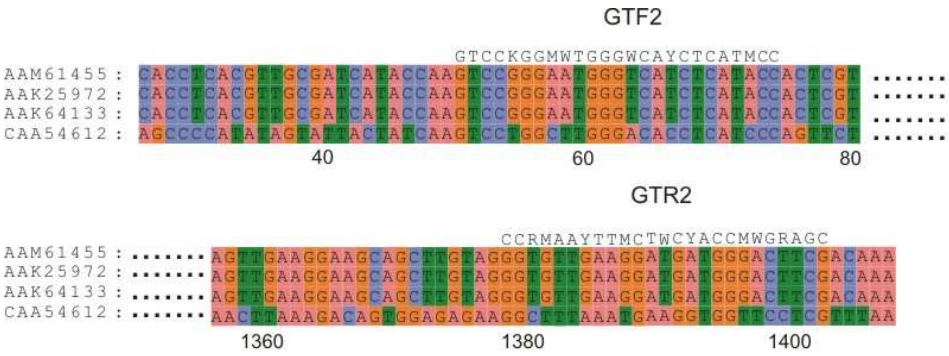
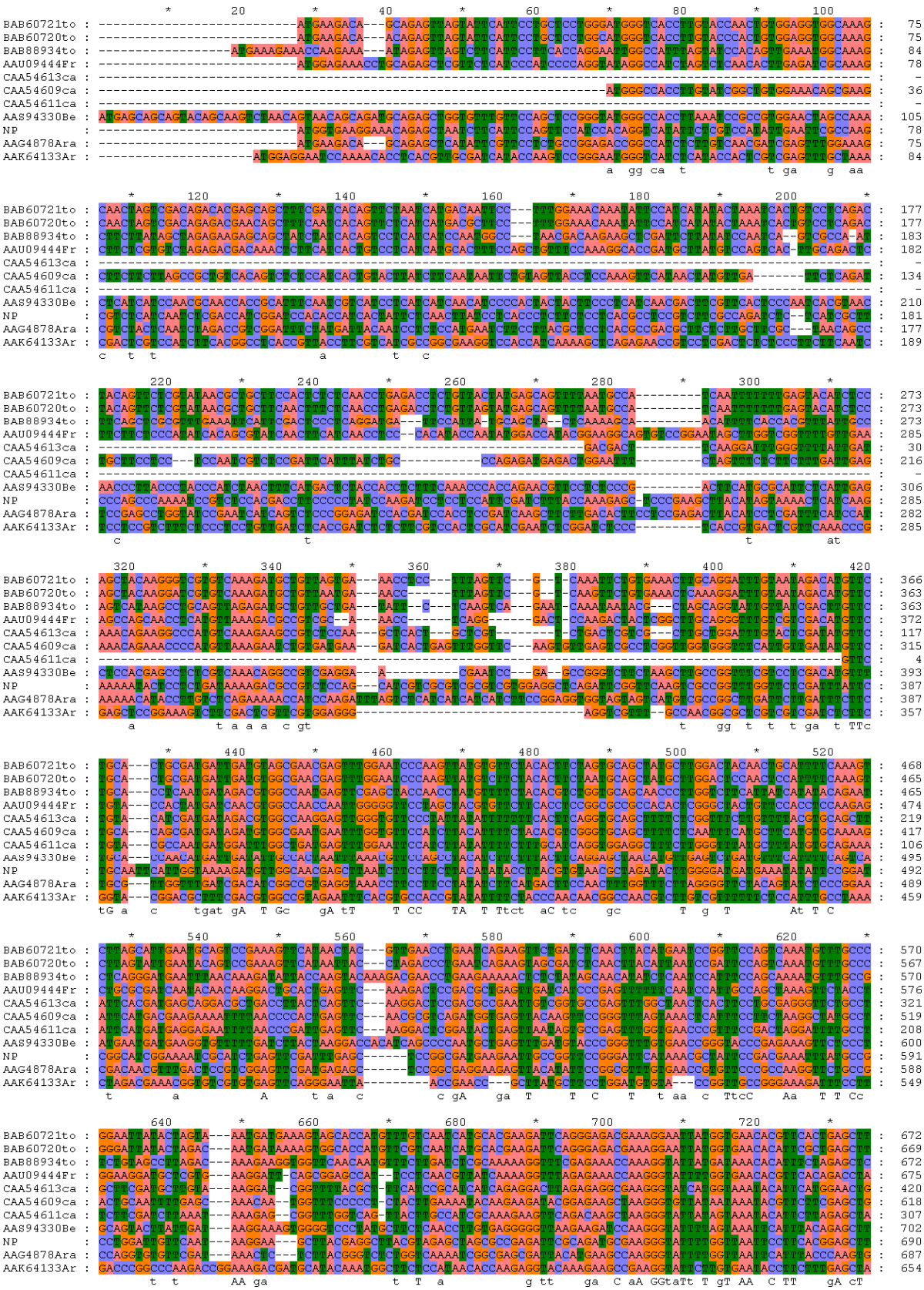
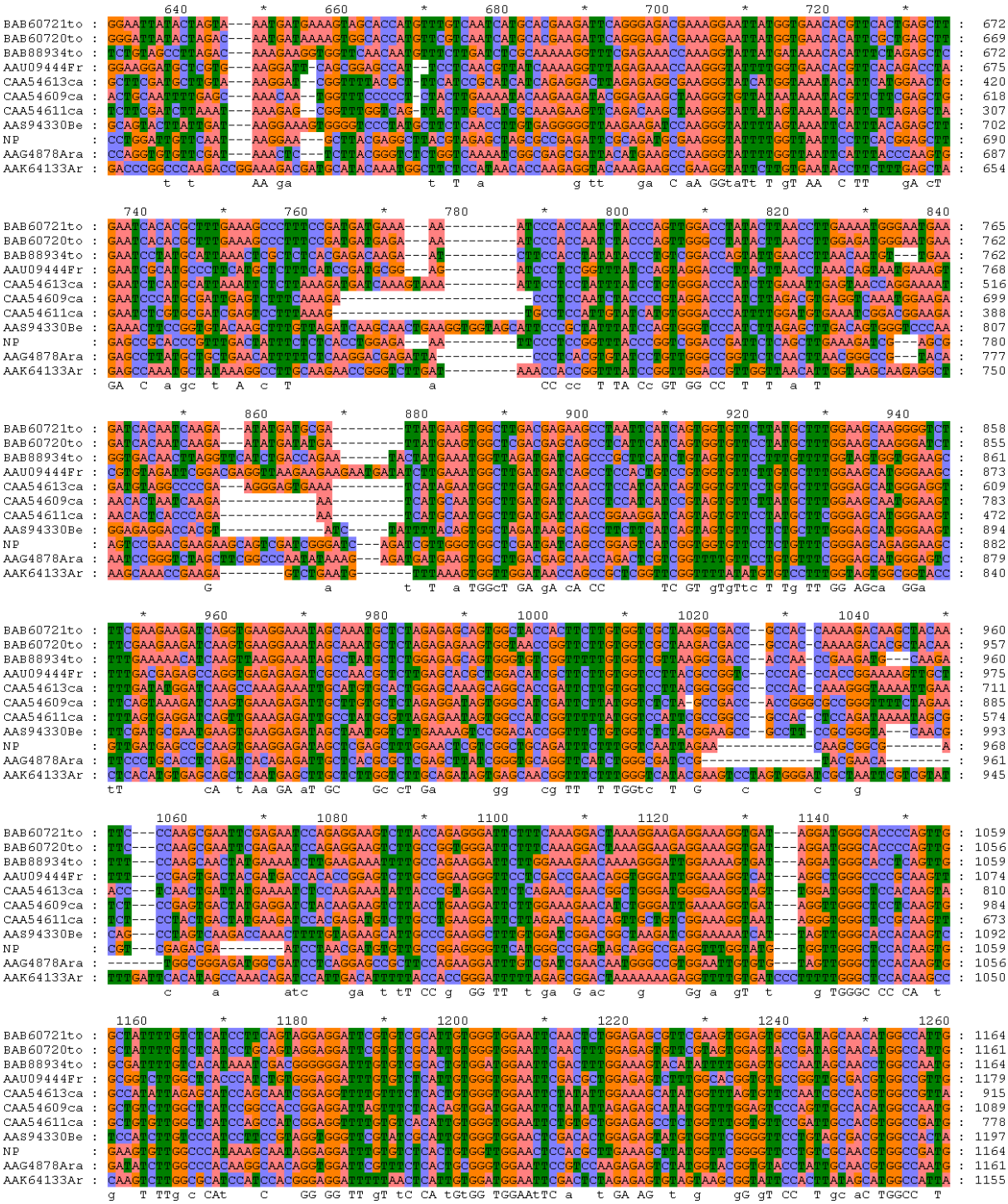


Fig. 5.13. Nucleotide alignment of isolates of group 2 GT. Amplification primers (GTF2/R2) are indicated. 5', 3' and core sequences have been omitted for illustration purposes.





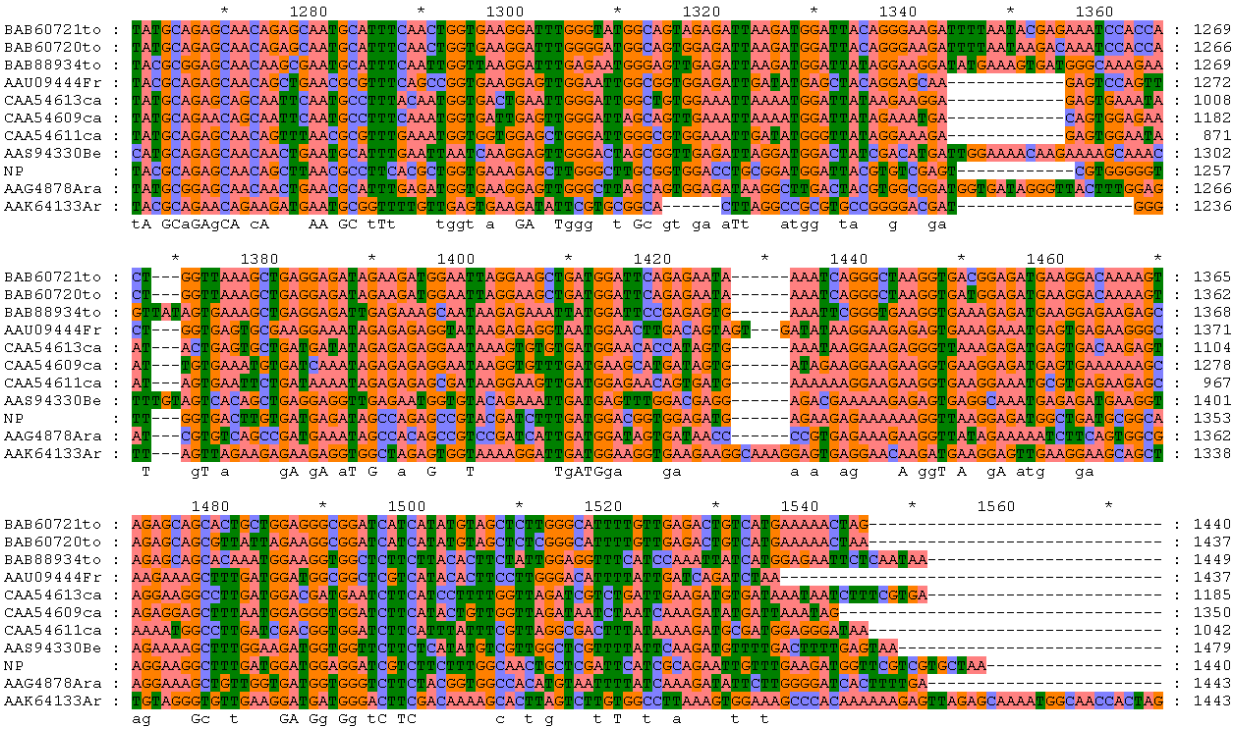


Fig. 5.14. Nucleotide alignment of group 3 glucosyltransferases.

In group 3 (Fig. 5.14) alignment, *A. thaliana* sequences NP_563784, AGG48783 and AAM65993 were omitted together with AAU09444 from *Fragaria x ananassa*, CAA54611 and CAA54613 from *M. esculenta* were omitted. The alignment was reperformed which resulted in more conserved regions to design forward and reverse primers GTF3 and R3 (Fig. 5.15).

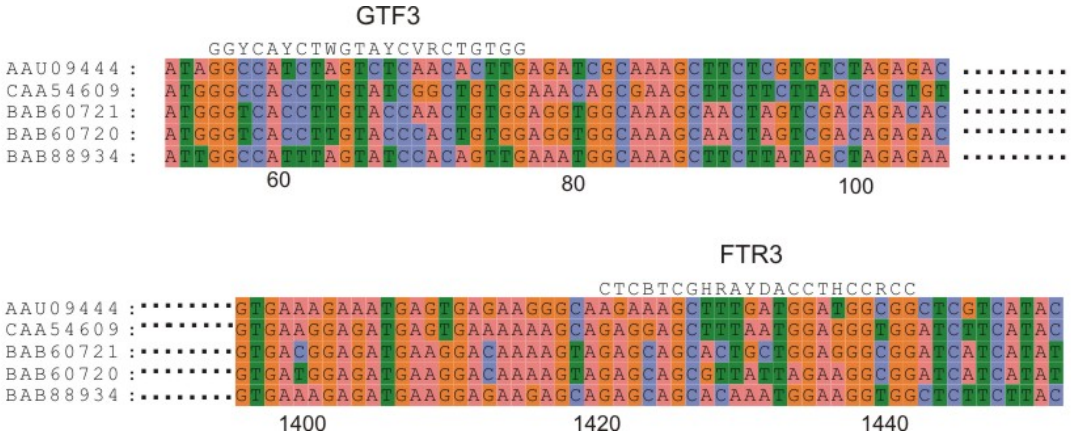
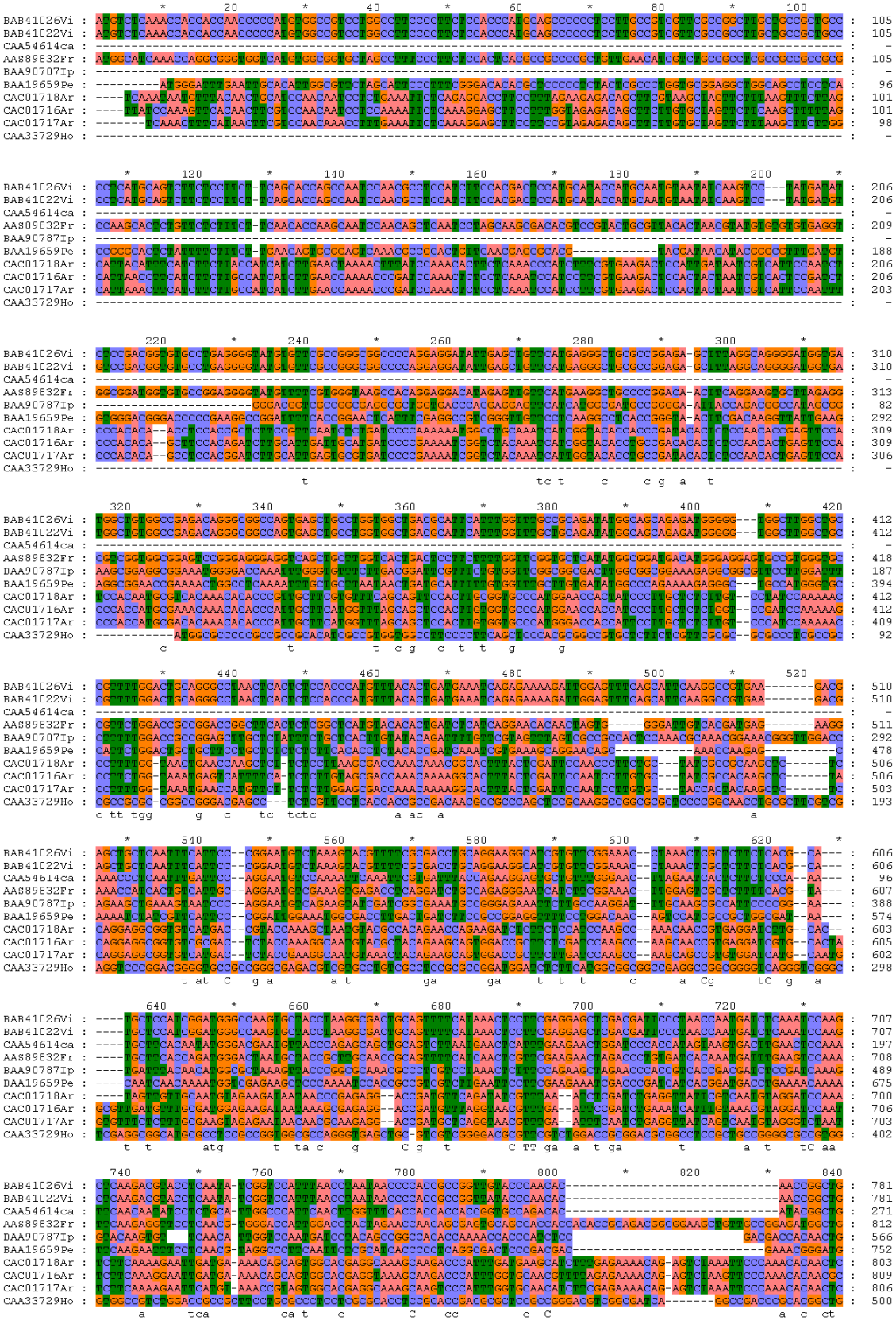


Fig. 5.15. Nucleotide alignment of isolates of group 3 GT. Amplification primers (GTF3/R3) are indicated. 5', 3' and core sequences have been omitted for illustration purposes.


BAB41026Vi
BAB41022Vi
CAA54614ca
AAS89832Fr
BAA90787Ip
BAA19659Pe
CAC01718Ar
CAC01716Ar
CAC01717Ar
CAA33729Ho

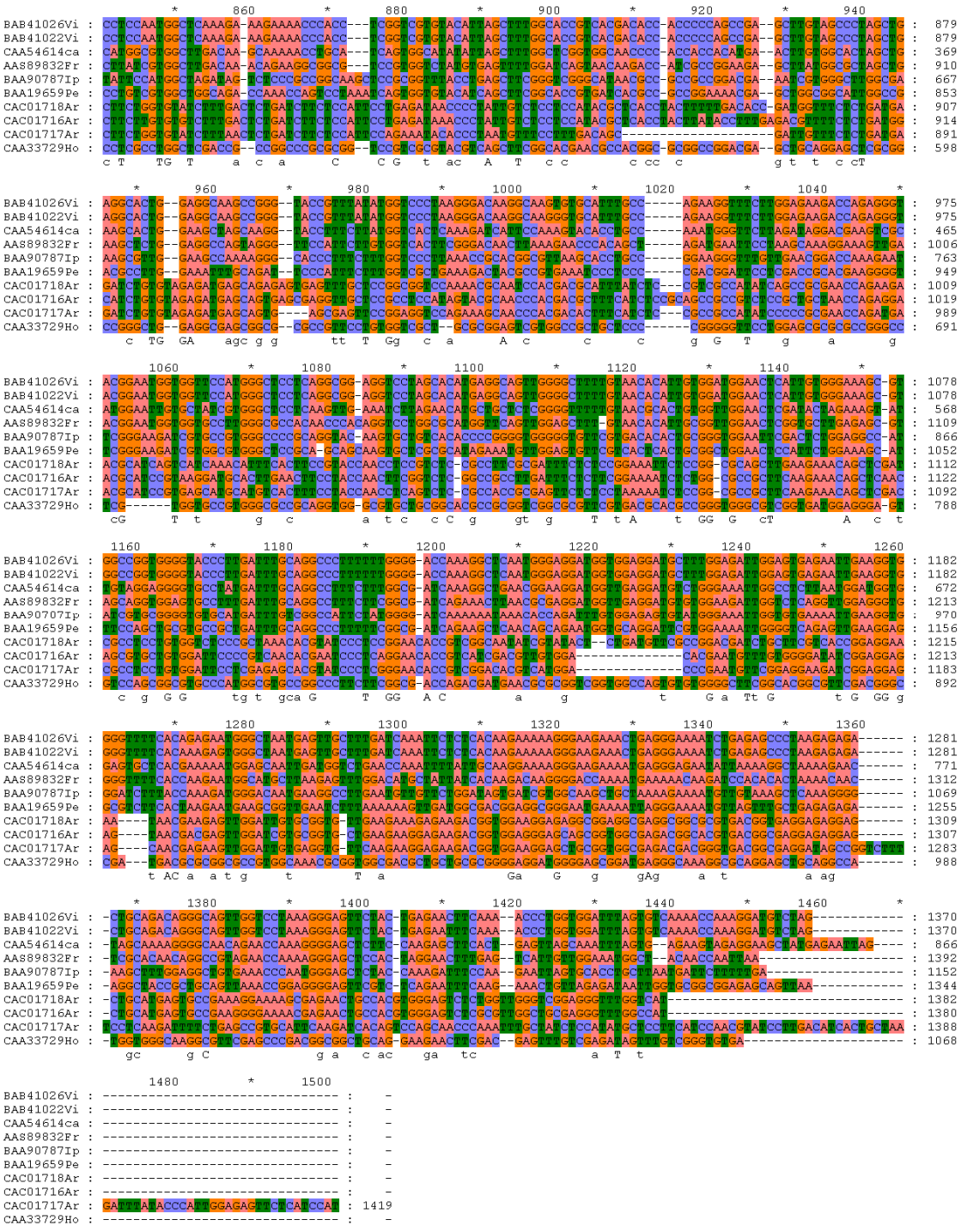


Fig. 5.16. Nucleotide alignment of group 4 glucosyltransferases.

Group 4 alignment (Fig. 5.16) showed few conserved regions so degenerate primers were difficult design. Some sequences (CAC01718, CAC01717 CAC01716 from *A. thaliana* , CAA33729 from *Hordeum vulgare* and BAA90787 from *Ipomoea batatas*) were omitted and the alignment repeated which resulted in more conserved regions from which to design forward and reverse primers GTF4 and R4 (Fig. 5.17).

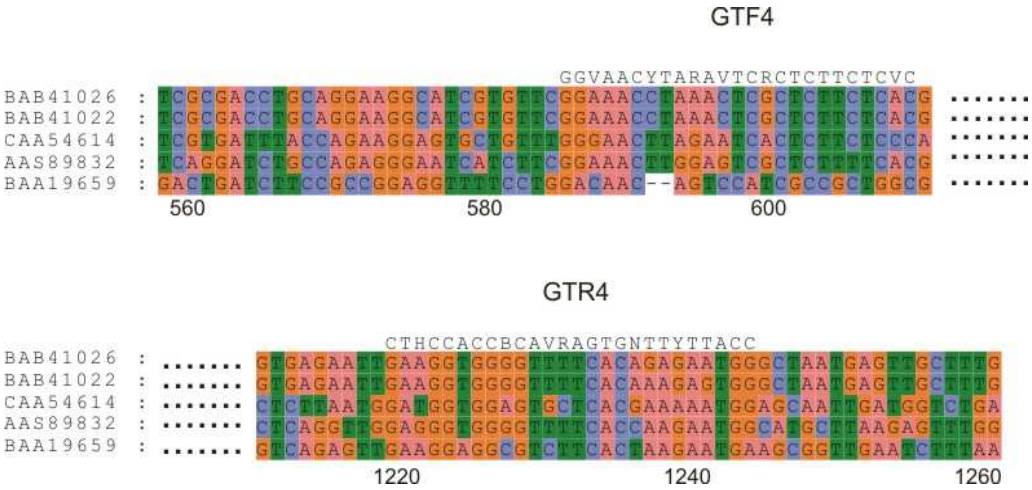


Fig. 5.17. Nucleotide alignment of isolates of group 4 GT. Amplification primers (GTF4/R4) are indicated. 5', 3' and core sequences have been omitted for illustration purposes.

Degenerate primers specific to these four selected groups of glucosyltransferases were designed (Table 5.9), and these primers were used to amplify glucosyltransferases from the cassava cDNA library.

Table 5.9

Designed forward and reverse degenerate PCR primers

Name	**Sequence (5'-3')**
GTF1	CWKTVWTGGCTCMTGGCCACATG
GTR1	GAWGAHCCWCCTYCTTCWAWWG
GTF2	GTCCKGGMWTGGGWCAYCTCATMCC
GTR2	CGARGWMCCAYCWTCMTTYAAMRCC
GTF3	CCRCCHTCCADYARHGCTBCTC
GTR3	GGYCAYCTWGTAYCVRCTGTGG
GTF4	GGVAACYTARAVTCRCTCTTCTCVC
GTR4	CCATTYTTNGTGARVACBCCACCHTC

Isolation and characterisation of cassava glucosyltransferases

A cDNA library was constructed to mRNA from deteriorated cassava roots (Reilly et al., 2007). Designed degenerate primers were used to amplify glucosyltransferases from the constructed cDNA library by PCR. For primers GTF2, GTR2 annealing temperature was at 52 ºC while for GTF4 and GTR4 it was at 55 ºC and for GTF1, GTR1, GTF3 and GTR3 temperatures between 45-55 ºC were used.

The PCR product was gel isolated, ligated into the TOPO® vector. then transformed into One Shot® Top10 Chemically competent *E. coli* cells using a TOPO TA Cloning® Kit. The plasmid DNA was purified and the cloned DNA sequenced using forward and reverse M13 universal primers. The following two primers; UNIF and UNIR were designed (Fig. 5.18) from the sequenced part to complete the sequencing in both directions.

UNIF: $^{5'}$CTTCAGAAGGAGCCCATG$^{3'}$

UNIR: $^{5'}$CTGCTGGCTCCGCTCAAG$^{3'}$

Primers GTF2 and GTR2 gave a PCR product (GT 2, 1.4 kb) (Fig. 5.18) that was 99% similar to CAA54612 (Fig. 5.19), a flavonoidal glucosyltransferase that had been previously isolated from cassava cotyledons (Hughes and Hughes, 1994). This sequence, CAA54612 was in the sequence alignment used to design GTF2 and GTR2 primers. This PCR product demonstrates the specificity of the designed primers.

```
   1  GAATTCGCCC TTGGTCCTGG CTTGGGTCAT CTCATCCCAG TTCTCGAACT
  51  GGGAAAACGC ATAGTTACAC TCTGCAACTT CGATGTGACT ATATTCATGG
 101  TGGGTTCCGA CACATCAGCT GCTGAACCTC AAGTTCTCCG ATCAGCCATG
 151  ACTCCAAAAC TCTGCGAAAT CATCCAACTC CCACCACCTA ACATTTCCTG
 201  CCTGATCGAC CCAGAAGCCA CCGTATGTAC CCGTCTTTTT GTTTTGATGA
 251  GAGAAATCAG GCCAGCTTTC CGGGCGGCAG TATCCGCTCT CAAGTTTCGA
 301  CCGGCAGCCA TAATTGTCGA CCTCTTTGGA ACTGAATCTC TGGAGGTAGC
 351  TAAAGAACTT GGCATCGCAA AATATGTGTA CATAGCTTCT AATGCATGGT
 401  TTTTAGCTCT TACTATATAT GTACCGATTC TAGATAAAGA GGTGGAAGGA
 451  GAGTTTGTTC TTCAGAAGGA GCCCATGAAA ATTCCTGGTT GCAGGCCGGT
 501  TCGACCGGAA GAGGTAGTCG ATCCCATGCT GGACCGAACC AATCAACAAT
 551  ATTCCGAGTA TTTTCGCTTA GGTATCGAGA TCCCGACAGC TGACGGTATA
 601  TTAATGAACA CGTGGGAAGC TCTTGAACCA ACAACATTCG GAGCTTTGAG
 651  AGATGTGAAG TTCCTGGGCC GAGTAGCTAA GGTACCGGTT TTTCCGATTG
 701  GTCCTCTGAG GAGACAGGCC GGACCGTGCG GTTCAAATTG TGAGTTACTT
 751  CGATTGGTTA GACCAACAAC CCAAAGAGTC GGTGGTTTAT GTGTCGTTTG
 801  GGAGTGGTGG GACTCTGTCA TTGGAGCAAA TGATCGAGCT TGCTTGGGGC
 851  CTTGAGCGGA GCCAGCAGAG GTTTATTTGG GTGGTTCGCC AACCGACCGT
 901  AAAGACAGGA GATGCAGCAT TTTTTACTCA AGGGGACGGT GCAGATGACA
 951  TGTCAGGGTA CTTCCCTGAG GGGTTCCTGA CCAGGATTCA GAACGTGGGG
1001  TTGGTGGTCC CACAATGGAG CCCACAAATT CACATCATGA GCCATCCATC
1051  AGTGGGAGTA TTTTTATCAC ACTGTGGTTG GAATTCTGTA TTGGAGAGCA
1101  TCACAGCAGG AGTGCCCATT ATTGCGTGGC CAATATATGC TGAGCAGAGG
1151  ATGAATGCGA CGCTGTTGAC GGAGGAGCTA GGCGTTGCAG TGAGGCCAAA
1201  GAATTTACCG GCGAAAGAAG TAGTGAAGAG GGAGGAGATA GAGAGGATGA
1251  TTAGAAGGAT TATGGTAGAT GAAGAAGGAA GTGAAATAAG AAAGAGAGTG
1301  AGAGAACTTA AAGACAGTGG AGAGAAGGCT TTAAATGATG AT
```

Fig. 5.18. Nucleotide sequence ofGTF2 and GTR2 PCR product (GT 2, 1.4kbp). Sequences used for designing sequencing primers UNIF and UNIR are underlined

```
Query  14    GTCCTGGCTTGGGTCATCTCATCCCAGTTCTCGAACTGGGAAAACGCATAGTTACACTCT  73
             |||||||||||||  || ||||||||||||||||||||||||||||||||||||||||||
Sbjct  81    GTCCTGGCTTGGGACACCTCATCCCAGTTCTCGAACTGGGAAAACGCATAGTTACACTCT  140

Query  74    GCAACTTCGATGTGACTATATTCATGGTGGGTTCCGACACATCAGCTGCTGAACCTCAAG  133
             |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Sbjct  141   GCAACTTCGATGTGACTATATTCATGGTGGGTTCCGACACATCAGCCGCTGAACCTCAAG  200

Query  134   TTCTCCGATCAGCCATGACTCCAAAACTCTGCGAAATCATCCAACTCCCACCACCTAACA  193
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  201   TTCTCCGATCAGCCATGACTCCAAAACTCTGCGAAATCATCCAACTCCCACCACCTAACA  260

Query  194   TTTCCTGCCTGATCGACCCAGAAGCCACCGTATGTACCCGTCTTTTTGTTTTGATGAGAG  253
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  261   TTTCCTGCCTGATCGACCCAGAAGCCACCGTATGTACCCGTCTTTTTGTTTTGATGAGAG  320

Query  254   AAATCAGGCCAGCTTTCCGGGCGGCAGTATCCGCTCTCAAGTTTCGACCGGCAGCCATAA  313
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  321   AAATCAGGCCAGCTTTCCGGGCGGCAGTATCCGCTCTCAAGTTTCGACCGGCAGCCATAA  380

Query  314   TTGTCGACCTCTTTGGAACTGAATCTCTGGAGGTAGCTAAAGAACTTGGCATCGCAAAAT  373
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  381   TTGTCGACCTCTTTGGAACTGAATCTCTGGAGGTAGCTAAAGAACTTGGCATCGCAAAAT  440

Query  374   ATGTGTACATAGCTTCTAATGCATGGTTTTTAGCTCTTACTATATATGTACCGATTCTAG  433
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  441   ATGTGTACATAGCTTCTAATGCATGGTTTTTAGCTCTTACTATATATGTACCGATTCTAG  500

Query  434   ATAAAGAGGTGGAAGGAGAGTTTGTTCTTCAGAAGGAGCCCATGAAAATTCCTGGTTGCA  493
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  501   ATAAAGAGGTGGAAGGAGAGTTTGTTCTTCAGAAGGAGCCCATGAAAATTCCTGGTTGCA  560

Query  494   GGCCGGTTCG-ACCGGAAGAGGTAGTCGATCCCATGCTGGACCGAACCAATCAACAATAT  552
             |||||||||| |||| ||||||||||||||| ||||||||||||| ||||||||||||
Sbjct  561   GGCCGGTTCGGACCG-AAGAGGTAGTCGATCCTATGCTGGACCGAACAAATCAACAATAT  619

Query  553   TCCGAGTATTTTCGCTTAGGTATCGAGATCCCGACAGCTGACGGTATATTAATGAACACG  612
             |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Sbjct  620   TCCGAGTATTTTCGCTTAGGTATCGAGATCCCAACAGCTGACGGTATATTAATGAACACG  679

Query  613   TGGGAAGCTCTTGAACCAACAACATTCGGAGCTTTGAGAGATGTGAAGTTCCTGGGCCGA  672
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  680   TGGGAAGCTCTTGAACCAACAACATTCGGAGCTTTGAGAGATGTGAAGTTCCTGGGCCGA  739

Query  673   GTAGCTAAGGTACCGGTTTTTCCGATTGGTCCTCTGAGGAGACAGGCCGGACCGTGCGGT  732
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  740   GTAGCTAAGGTACCGGTTTTTCCGATTGGTCCTCTGAGGAGACAGGCCGGACCGTGCGGT  799

Query  733   TCAAATTGTGAGTTACTTCGATTGGTTAGACCAACAACCCAAAGAGTCGGTGGTTTATGT  792
             |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Sbjct  800   TCAAATTGTGAGTTACT-CGATTGGTTAGACCAACAACCCAAAGAGTCGGTGGTTTATGT  858

Query  793   GTCGTTTGGGAGTGGTGGGACTCTGTCATTGGAGCAAATGATCGAGCTTGCTTGGGGCCT  852
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  859   GTCGTTTGGGAGTGGTGGGACTCTGTCATTGGAGCAAATGATCGAGCTTGCTTGGGGCCT  918

Query  853   TGAGCGGAGCCAGCAGAGGTTTATTTGGGTGGTTCGCCAACCGACCGTAAAGACAGGAGA  912
             |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Sbjct  919   TGAGCGGAGCCAGCAGAGGTTTATTTGGGTGGTTCGCCAACCCACCGTAAAGACAGGAGA  978


Query  913   TGCAGCATTTTTTACTCAAGGGGACGGTGCAGATGACATGTCAGGGTACTTCCCTGAGGG  972
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Sbjct  979   TGCAGCATTTTTTACTCAAGGGGACGGTGCAGATGACATGTCAGGGTACTTCCCTGAGGG  1038

Query  973   GTTCCTGACCAGGATTCAGAACGTGGGGTTGGTGGTCCCACAATGGAGCCCACAAATTCA  1032
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Sbjct  1039  GTTCCTGACCAGGATTCAGAACGTGGGGTTGGTGGTCCCACAATGGAGCCCACAAATCCA  1098

Query  1033  CATCATGAGCCATCCATCAGTGGGAGTATTTTTATCACACTGTGGTTGGAATTCTGTATT  1092
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1099  CATCATGAGCCATCCATCAGTGGGAGTATTTTTATCACACTGTGGTTGGAATTCTGTATT  1158

Query  1093  GGAGAGCATCACAGCAGGAGTGCCCATTATTGCGTGGCCAATATATGCTGAGCAGAGGAT  1152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1159  GGAGAGCATCACAGCAGGAGTGCCCATTATTGCGTGGCCAATATATGCTGAGCAGAGGAT  1218

Query  1153  GAATGCGACGCTGTTGACGGAGGAGCTAGGCGTTGCAGTGAGGCCAAAGAATTTACCGGC  1212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1219  GAATGCGACGCTGTTGACGGAGGAGCTAGGCGTTGCAGTGAGGCCAAAGAATTTACCGGC  1278

Query  1213  GAAAGAAGTAGTGAAGAGGGAGGAGATAGAGAGGATGATTAGAAGGATTATGGTAGATGA  1272
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1279  GAAAGAAGTAGTGAAGAGGGAGGAGATAGAGAGGATGATTAGAAGGATTATGGTAGATGA  1338

Query  1273  AGAAGGAAGTGAAATAAGAAAGAGAGTGAGAGAACTTAAAGACAGTGGAGAGAAGGCTTT  1332
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1339  AGAAGGAAGTGAAATAAGAAAGAGAGTGAGAGAACTTAAAGACAGTGGAGAGAAGGCTTT  1398

Query  1333  AAATGA  1338
              ||||||
Sbjct  1399  AAATGA  1404
```

Fig. 5.19. Comparison of the GT 2 sequence with the cassava cotyledon glucosyltransferase CAA54612 showing 99% identity.

The GTF4 and GTR4 PCR product (GT 4, 0.6 kb) (Fig. 5.20) was sequenced in both direction using forward and reverse universal primers M13. It showed 99% similar to CAA54614, which, is another flavonoidal glucosyltransferase that had previously been identified from cassava cotyledons (Hughes and Hughes, 1994). With GTF1&3 and GTR1&3, the PCR product was of insufficient quantity or quality for further investigation (no PCR product or the product size was not of the expected size ~1 kb).

Trials using different forward and reverse primer combinations were also carried out, e.g. using GTF1 with GTR2 and so on, at different annealing temperatures (45-55 ºC), but these trials were unsuccessful, suggesting that GTF1&3 and GTR1&3 either show excessively high degeneracy to anneal successfully to the target DNA under the wide range of conditions attempted or that cDNAs for glucosyltransferases of these classes were not represented in the PPD-related cDNA library used as the source of target.

```
  1  ACCATCCATT AAGAGGCCAA TTTCCCAGAC ATCCTCAACC ATCCTTCCGT
 51  TCAGCCTTTG ATCGCCAAAG AAAGGCCTGC AAATCATAGG CACCCCTCCT
101  ACAATACTTT CTAGTATCGA GTTCCAACCA CAGTGCGTTA CAAAGACCCC
151  GAGAGCAGCA TGTTCTAAGA TTTCAACTTG AGGAGCCCAC GATAGCACAA
201  TTCCATGCGA CTTCGTCCTA TCTAAGAACC CATTTGGCAG GTGTACTTTG
251  GAATGATCTT TGAGTGACCA TAAGAAAGGT ACCTTGCTAG CTTCCAGTGC
301  TTCAGCTAGT GCCACAAGTT CATGTGGTGG TGGGGTTGCC ACCGAGCCAA
351  AGCTAATATA TGCCACTGAT GCAGGTTTTT GCTTGTCAAG CCACGCCATG
401  CAGCCGTATG TGTCTGGCAC CGGTGGTGGT GGTGAAACCA AGTTGAATGG
451  GCCAATGCAG AGGATATTGT TGAATTTGGA GTTCAAGTCA CTTACTATGG
501  TGGGATCCAG TTCTTCAAAT GAGTTCATTA AGACTGCAGC TGCTCTGGGT
551  AACATTCGTC CCATATTGTG AAGCATTTGC GAGAAGAGCG ACTCTAAGTT
601  CCCAA
```

Fig. 5.20. GTF4 and GTR4 PCR product (GT 4, 605 b) was 99% similar to CAA54614 which is another cassava cotyledon glucosyltransferase.

The two sequences successfully cloned and identified probably represent flavonoidal glucosyltransferases (sequence similarity was identified by BLAST searching) from the cDNA library of deteriorated cassava root that are expressed during PPD and also expressed in cassava cotyledons (cassava young seedlings, CM1223-11 (Hughes and Hughes, 1994)) and could be involved in the glucosylation of scopoletin into scopolin and esculetin into esculin during PPD. This is the first time that these two genes have been identified as expressed in cassava roots during PPD. To my knowledge there are no reports of the presence of flavonoidal glycosides from fresh or deteriorated cassava root (cf kaempferol-3-*O*-rutinoside and rutin from leaves); however, the structural similarity between flavonoids and coumarins supports the suggestion that these enzymes could be involved in coumarin glucosylation in deteriorating cassava roots.

In order to complement and extend the range of glucosyltransferases known to be expressed during PPD, the PPD-related cDNA library was screened by hybridising plaque lifts with radio-labelled gene-probes for glucosyltransferases with established biological functions in other plants. These probes were: tobacco scopoletin glucosyltransferase, anthocyanin glucosyltransferase from *Perilla frutescens* (purple mint, sometimes wild red basil) and cinnamate glucosyltransferase from *Fragaria* x ananassa (strawberry) (Fraissinet-Tachet et al., 1998; Yamazaki et al., 1999; Lunkenbein et al., 2006). *Perilla frutescens* anthocyanin glucosyltransferase (accession number, AB 013596) was chosen as a probe because the corresponding enzyme added a glucose unit to anthocyanin which is structure similarity to hydroxycoumarin. The gene has also been expressed in a yeast vector and protein activity was analysed using UDP-glucose and cyaniding 3-*O*-glucoside forming cyaniidin3,5-*O*-diglucoside. The gene was given as a gift from Prof. Mami Yamazaki (Chiba University, Japan) (Yamazaki et al., 1999). *Fragaria* cinnamate glucosyltransferase (FaGT2 accession number, AY663785) was chosen as a probe because of added a glucose unit to cinnamate and its derivatives which have structure similarity to hydroxycoumarin. Assays with FaGT2 show that cinnamic acid, benzoic acid and their derivatives were accepted as substrates. FaGT2 was a gift from Prof. Wilfried Schwab (Technical University Munich, Germany) (Lunkenbein et al., 2006). Tobacco scopoletin glucosyltransferase (TOGT1 accession number, AAK28303) was chosen as a probe because of its ability to add a glucose unit to scopoletin. TOGT1 catalytic activity has been tested by its expression in *E. coli*. The enzyme catalyzed the transfer of the glucosyl moiety onto the carboxyl and the hydroxyl groups. Among the phenylpropanoids tested, scopoletin and esculetin were the best substrates, followed by cinnamic and the hydroxycinnamic acids, *p*-coumaric, *o*-coumaric, caffeic and ferulic acids. TOGT1 was amplified from both cDNA and genomic DNA of tobacco and there were no differences in the sequences, which indicated that this gene did not contain any intron (Fraissinet-Tachet et al., 1998).

Tobacco scopoletin glucosyltransferase coding sequence was amplified by PCR from tobacco genomic DNA using the following TOGTF1 and TOGTR1 primers. .:
Forward primer (TOGTF1): $^{5'}$CGGATCCATGGGTCAGCTCCATWTTTTC$^{3'}$
Reverse primer (TOGTR1): $^{5'}$CGGATCCTTAATGACCAGTAGAACTATATG$^{3'}$

These primers have been used before to amplify the glucosyltransferase cDNAs TOGT1 and TOGT2 (accession numbers U32644 and U32643, respectively). TOGT1 and TOGT2 are glucosyltransferases which revealed the significant catalysis of the transfer of UDP-glucose to scopoletin and esculetin forming the corresponding glucoside, scopolin and esculin (Fraissinet-Tachet et al., 1998).

PCR sample was performed (annealing temperature was 45 ºC for 1 min). The PCR product (1.4 kb) (Figs. 5.21) was gel isolated, cloned then transformed to competent *E. coli* cells. The plasmid DNA was purified and sequenced. The 1.4 kb fragment showed 100% similarity to TOGT1. The purified plasmid was digested with *Bam*HI (NEB, # R0136S) (Fig. 5.22) to use the gene fragment as a probe to screen the PPD-related cDNA library.

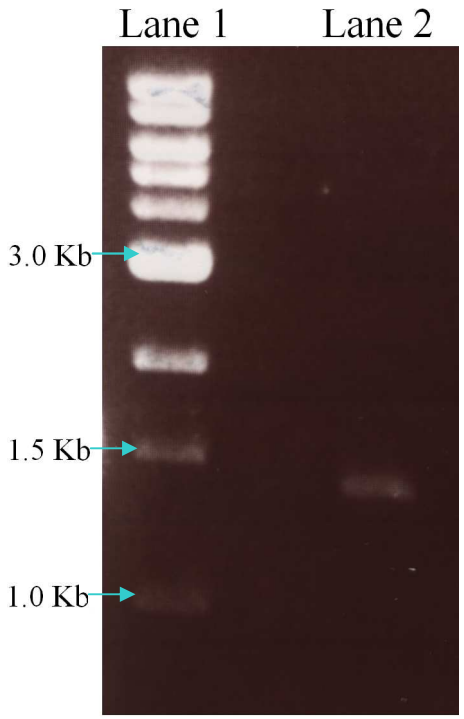


Fig. 5.21. PCR amplification of TOGT1 glucosyltransferase from tobacco genomic DNA. PCR product (lane 2) was run on an ethidium bromide stained 1.0% agarose gel. The size marker (lane 1) is Biolab DNA 1.0 kb ladder.

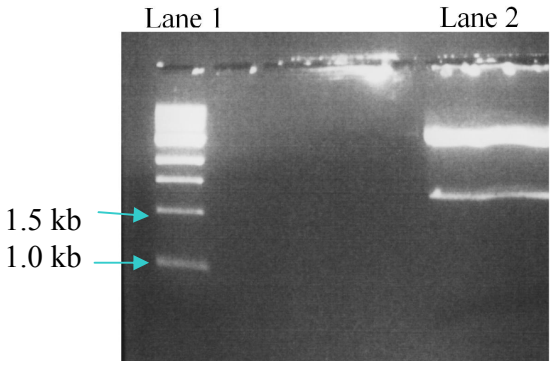


Fig. 5.22. Restriction enzyme digestion using *Bam* HI. The digestion product (lane 2) was run on ethidium bromide stained 1.0% agarose gel. The size marker (lane 1) is Biolab DNA 1.0 kb ladder.

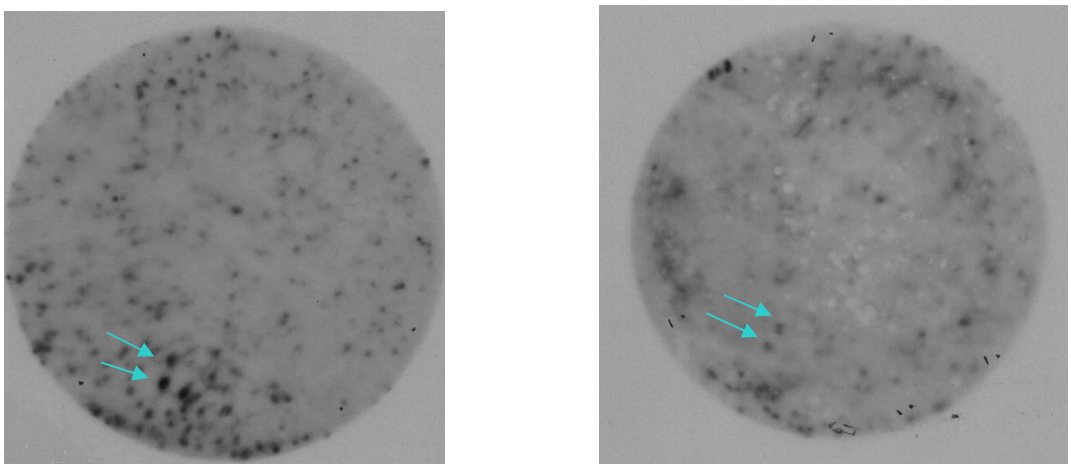


Fig. 5.23. Typical duplicate autoradiographs from plaque screening of cassava cDNA library after post harvest physiological deterioration. The arrows indicate duplicate positive plaques.

Screening the cassava PPD-related cDNA library using the labelled TOGT1 probe led to the identification of two positive plaques (Fig. 5.23). These plaques were purified, the corresponding plasmids excised and their insert DNAs sequenced in both directions using M13 forward and reverse primers from Bluescript plasmids after their excision and purification (Fig. 5.24).

```
       1 ACTGTCACACCCTGCAATCGGAGGTTTCTTAACACATTGTGGTTGGAATTCAACACTAGA
b        L  S  H  P  A  I  G  G  F  L  T  H  C  G  W  N  S  T  L  E  -

      61 AGCAATAAGTAGTGGCGTGCCTGTAATCACATGGCCACTCTTTGCAGACCAATTTTGTAA
b        A  I  S  S  G  V  P  V  I  T  W  P  L  F  A  D  Q  F  C  N  -

     121 TGAGAAGTTAGCAGTGCAGGTACTGAAGATAGGCGTCAGAGTTGGAGTGGAGGTTCCTGA
b        E  K  L  A  V  Q  V  L  K  I  G  V  R  V  G  V  E  V  P  E  -

     181 GAGGGGAGTGGAGGAAGGGCAGAATGGTGTATTGGTGAAAAAGGAAGATGTTGAGAAGGC
b        R  G  V  E  E  G  Q  N  G  V  L  V  K  K  E  D  V  E  K  A  -

     241 GTTGAATGATTTGATGAAAGAAGGAGAGGAAAGGGAAGAGAGAAGAAGAAGAGTTAAAGA
b        L  N  D  L  M  K  E  G  E  E  R  E  E  R  R  R  R  V  K  E  -

     301 GCTTGCAGAGATGGCAAAGAAAGCAACTGAGGAAGGAGGTTCTTCTTGTCTCAACATCAA
b        L  A  E  M  A  K  K  A  T  E  E  G  G  S  S  C  L  N  I  K  -

     361 ACTTCTAATCCAAGATATCATGCAGCACGTAAATGAAGAAGCATCGCCTCCTCCAATTAA
b        L  L  I  Q  D  I  M  Q  H  V  N  E  E  A  S  P  P  P  I  K  -

     421 ATATCCAGAAAATTCCTAAAGTCCCTAGAAAAATGTGGTGCTCTCTTCAAGAAGCTTCAA
b        Y  P  E  N  S  *

     481 CAAGTTTCTTCATCCGTGCAATTCTCATCACCGGCGTATGCAGAAACTACTGTACCAACT

     541 GCTAGTGAGAATGAAGGGAACACTCGAGTGTGTGAGTGGATTGTGTTTGAGTGGTGACAA

     601 TGTGTGCACGTGAAAACATTTGTATACAATCGGTTACCGTAACAGCCCGAAACCGAACTG

     661 CTACCGGCACTAGGATTCAGATCGATTTAAGGTCGCCGGAACCCGTAGTAAGTCTGCTAT

     721 CCTGTTTGTGTACCTGTGAAATCTCATACATGATCATTCTTGTTCTGTCAAACTATTAAA

     781 ACTGTTCTGATACAAAAAAAAAAAAAAAAA
```

Fig. 5.24. Nucleotide and amino acid sequence of novel glucosyltransferase (GT 8) identified using TOGT1 as a labelled probe. Consensus sequence is in bold.

A Blast search revealed that one of these positive plaques corresponding to the partial sequence (800 bp) from a novel gene from cassava (GT 8). Its predicted amino acid sequence show 65 % identity to flavonoidal glucosyltransferases from *Nicotiana tabacu*m (BAD93688.1) and from *Medicago truncatula* (ABI94020.1). Also it has some 48 % similarity to *Nicotiana* scopoletin glucosyltransferase (TOGT1) which had been used as the probe for cDNA-library screening. The relationships between GT 8 and other GTs can be seen from the tree of various glucosyltransferases from different plants (Fig. 5.25). The sequence from the other positive plaque did not show any similarity to glucosyltransferases.

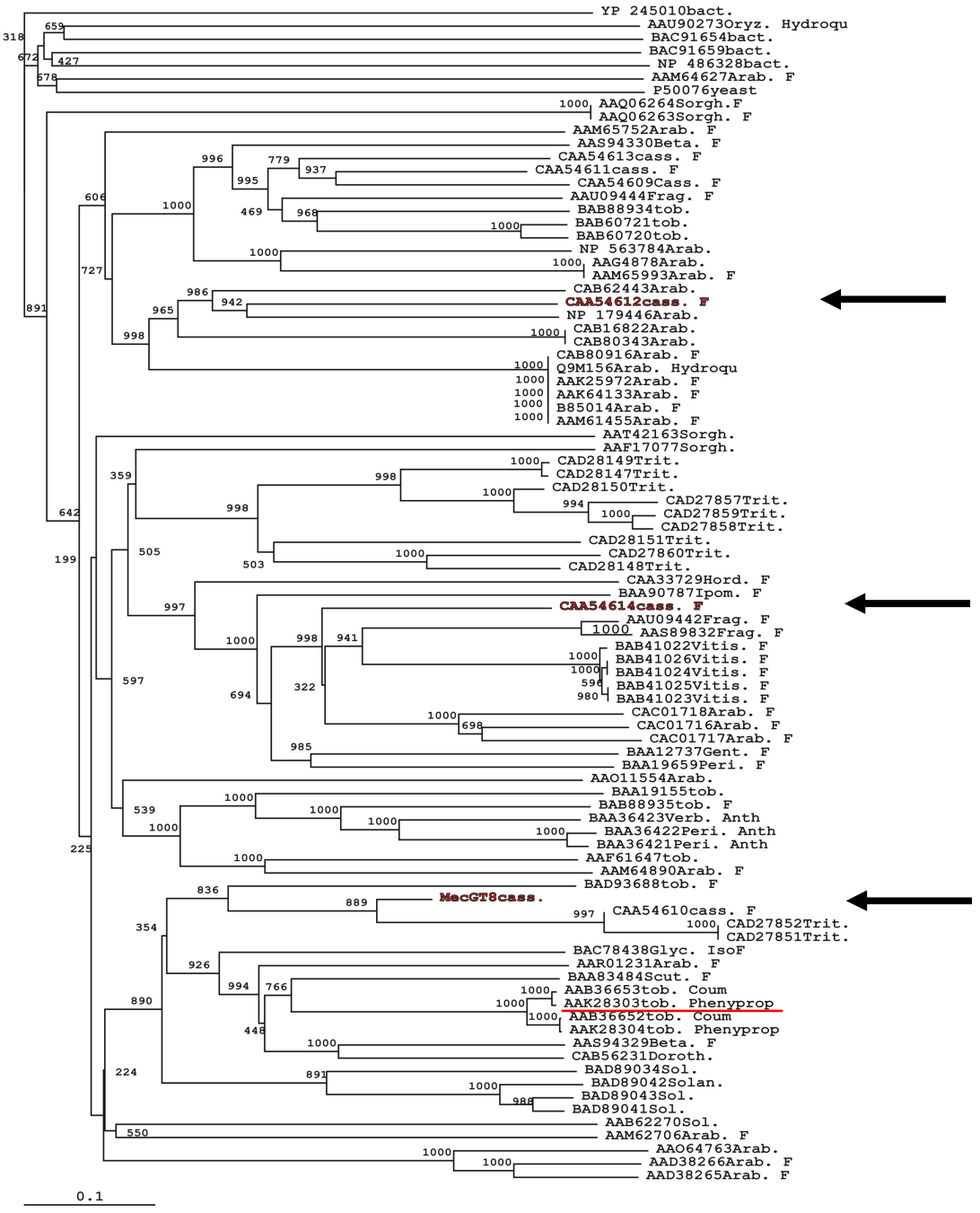


Fig. 5.25. Neighbour-Joining bootstrap tree (designed using Clustal X) illustrating sequence similarity relationships of different glucosyltransferases from various plants and the arrows referring to the newly identified ones. Abbreviations: Anthocyanin GT (Anth), Coumarin (Coum), Flavonoid (F), Hydroquinone (Hydroqu) and Phenylpropanoid (Phenylprop). TOGT1 which used as probe is underlined.

Based on sequence data, researchers have proposed a consensus sequence, called the PSPG box, for plant glycosyltransferases that are involved in secondary metabolism (Hughes and Hughes, 1994; Vogt and Jones, 2000). This sequence consists of about 40 amino acids in the C-terminal region of the proteins, and is thought to be the binding site for the nucleotide-diphosphate sugar. The predicted amino acid sequence from the partial cDNA GT8 includes only 37 of this 40 PSPG box; however, this is sufficient to confirm with reasonable certainty that the isolated clone corresponds to a glycosyltransferase (Fig. 5.26).

```
The consensus Nterm < WAPQVEVLAHPAVGCFV THCGWNSTLE SISAGVPMVA WPFFADQ >Cterm
GT8                          LSHPAIGGFL THCGWNSTLE AISSGVPVIT WPLFADQ
                            * *** * *  **********  ** ***    *******
```

Fig. 5.26. PSPG-box consensus sequence of plant secondary product glycosyltransferases. Highly conserved amino acids are shaded in blue (identity 50%) or red (identity 80%) reported in (Vogt and Jones, 2000) compared with the amino acid sequence of GT8.

Expression of identified glucosyltransferases in cassava organs

The isolation of three glucosyltransferases from a PPD-related cDNA library is strong evidence that these genes are expressed in the cassava storage root during PPD. However to confirm this and compare expression in the root and leaves, expression studies were performed using RT-PCR. Specific primers were designed for the three genes and used to amplify reverse-transcribed RNA isolated from freshly harvested roots (time 0), roots after 24 and 72 h after harvest, together with fresh cassava leaves. In addition, primers were used to amplify glyceraldehyde-3-phosphate dehydrogenase (GADPH) as a control. Primers designed from GADPH gene have been used as a control in RT-PCR in plants (Richert et al., 1996). GADPH (accession number, AF136128) was identified in cassava (Olsen and Schaal, 1999). GADPHF, a forward primer and GAPDHR, a reverse primer were designed from AF136128. The primers used are listed in Table 5.10.

Table 5.10

Designed forward and reverse degenerate PCR primers

Name	**Sequence (5'-3')**	**Gene amplified**
GTF21	CTC CGA TCA GCC ATG AC	GT2
GTR21	GTT CGG TCC AGC ATA GG	
GTF42	GAC GAA GTC GCA TGG AAT	GT4
GTR42	CCA GAC TCT TCA CCA TTC	
GTF84	GAA GTT AGC AGT GCA GGT	GT8
GTR83	CAG GTA CAC AAA CAG GAT AG	
GADPHF	CAG AAG ACT GTT GAC GGC CC	Glyceraldehyde-3-Phosphate
GAPDHR	CAA TTC CAG CCT TGG CGT C	Dehydrogenase

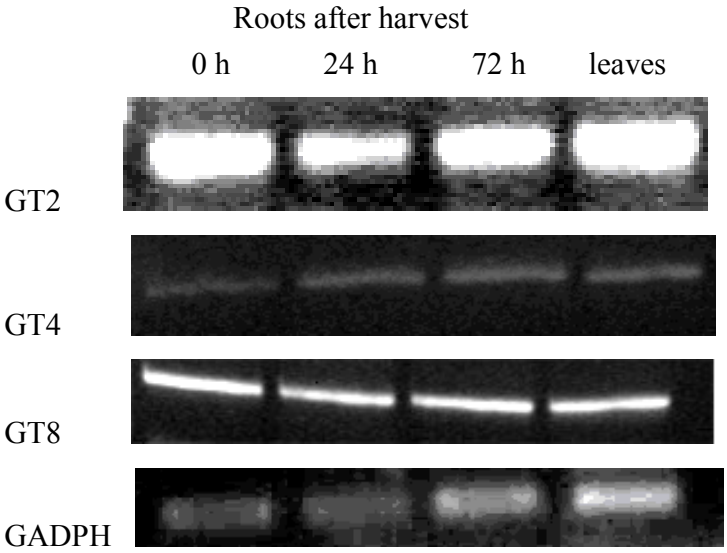


Fig. 5.27. Expression of three GTs and a control gene in cassava leaves and in harvested storage roots using RT-PCR.

All the three genes were expressed in the fresh and deteriorated roots as well as in fresh leaves (Fig. 5.27). GT2 expression slightly decreased after one day of root harvesting then increased after three days of root harvesting. GT4 may be expressed slightly less in fresh root and its expression increased along the deterioration period. GT8 is equally expressed in fresh and deteriorated cassava roots. However, no dramatic differences are apparent over the time-course of PPD or between roots and leaves. It is concluded that these genes may not play an important role specifically during PPD.

Chapter 6
General Conclusions and Future Work

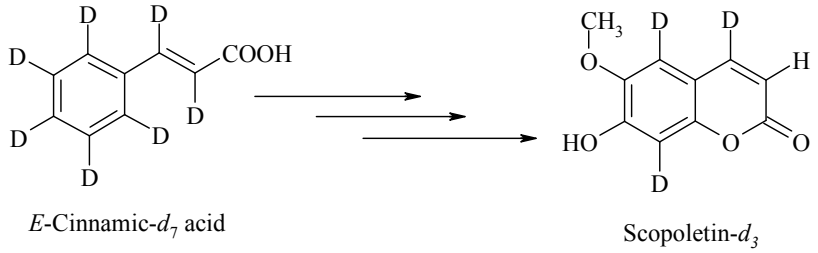
The tuberous root of cassava is the fourth most important food source in the tropics after rice, maize, and sugar cane; it is a staple food for more than 500 million people due to its high starch content. The cassava crop possesses an outstanding ability to grow in unfavourable environments. However, as an economic plant there are some problems that limit its utilization. Post-harvest physiological deterioration (PPD) is a major constraint for the development of this crop. Harvested roots have a short shelf-life of only one to three days due PPD becoming unpalatable and unmarketable and causing significant wastage and economic loss. PPD is explained as an oxidative burst which occurs within 15 min of the root being injured, followed by altered gene expression and the accumulation of secondary metabolites. Amongst these are hydroxycoumarins (e.g. scopoletin and its glucoside scopolin), which have different bioactivities and contribute essentially to the persistence of plants, being involved in processes such as defence against phytopathogens, response to abiotic stresses, regulation of oxidative stress, and probably hormonal regulation.

Hydroxycoumarins have different potential therapeutic uses due to antimicrobial, antioxidant, anti-inflammatory, anti-hepatotoxic, anticancer, hypotensive and spasmolytic activities. Hydroxycoumarins, as other phenylpropanoids are biosynthesised from L-phenylalanine, which is deaminated to cinnamate which is then *p*-hydroxylated to form *p*-coumarate. *p*-Coumarate can pass through three different pathways in different plants according to the order of its hydroxylation, methylation and lactonisation to synthesise scopoletin. Scopoletin is glucosylated by glucosyltransferase into its glucoside scopolin. Although hydroxycoumarins are important for plants and have many therapeutic actions, there are still some steps in their biosynthesis that are unresolved.

We exploit the accumulation of these coumarins during PPD to investigate the *E*-*Z*-isomerisation, *o*-hydroxylation, the proposed different pathways in their biosynthesis and glucosylation. The aim in understanding the biosynthesis of hydroxycoumarins in cassava roots is to aquire further understanding of the process of PPD. Ultimately, this may contribute towards the control of this problem and the production of a viable crop with long shelf life as a source for carbohydrates, not just for tropical regions, but for the world.

The ethanolic extracts of fresh and deteriorated cassava roots were analysed by chromatographic methods to isolate and identify the constituents (Table 2.4) by a variety of spectroscopic techniques (UV, NMR, MS).

The accumulation of hydroxycoumarins, mainly scopoletin and scopolin, during PPD was shown by TLC and HPLC. The incorporation of cinnamic acid-d_7 into hydroxycoumarins by cassava roots under PPD was investigated, in order to test the suitability of a simple chopped root system for biosynthetic studies. The unexpected experimental result (Fig. 6.1) that scopoletin-d_3 was produced, and not scopoletin-d_4, led us to study in detail the *E*-*Z*-isomerisation step in the biosynthesis of scopoletin and scopolin. This isomerisation is not resolved in plants but is the most likely step at which a proton loss might occur. It may be either a photochemical or enzymatic step, and has not been previously reported in cassava during PPD.



E-Cinnamic-d_7 acid

Scopoletin-d_3

Fig. 6.1. Feeding cassava roots with *E*-cinnamic-d_7 acid, scopoletin-d_3 was biosynthesised

Feeding *E*-cinnamic-2,3,2',3',4',5',6'-d_7 acid, *E*-cinnamic-3,2',3',4',5',6'-d_6 acid gave the same product which is scopoletin-d_3 acid and feeding with *E*-cinnamic-2-d_1 acid produced unlabelled scopoletin. These results confirmed that the *E*-*Z*-isomerisation of the C=C double bond in the biosynthesis of scopoletin and scopolin in cassava roots during PPD involves the specific exchange of the hydrogen atom at position 2 of *E*-cinnamic acid and thus is likely to be enzyme catalysed and not photochemical.

Three hypothetical pathways for the biosynthesis of scopoletin (Fig. 6.2) via: 2',4'-dihydroxycinnamate, 3',4'-dihydroxycinnamate (caffeate), or 4'-hydroxy-3'-methoxycinnamate (ferulate) have been proposed from studies in various plant species. Direct and competition feeding experiments with potential intermediates were carried out. We report feeding experiments in cassava roots under PPD with stable isotopically labelled *p*-coumaric acid, caffeic acid, ferulic acid, umbelliferone and esculetin.

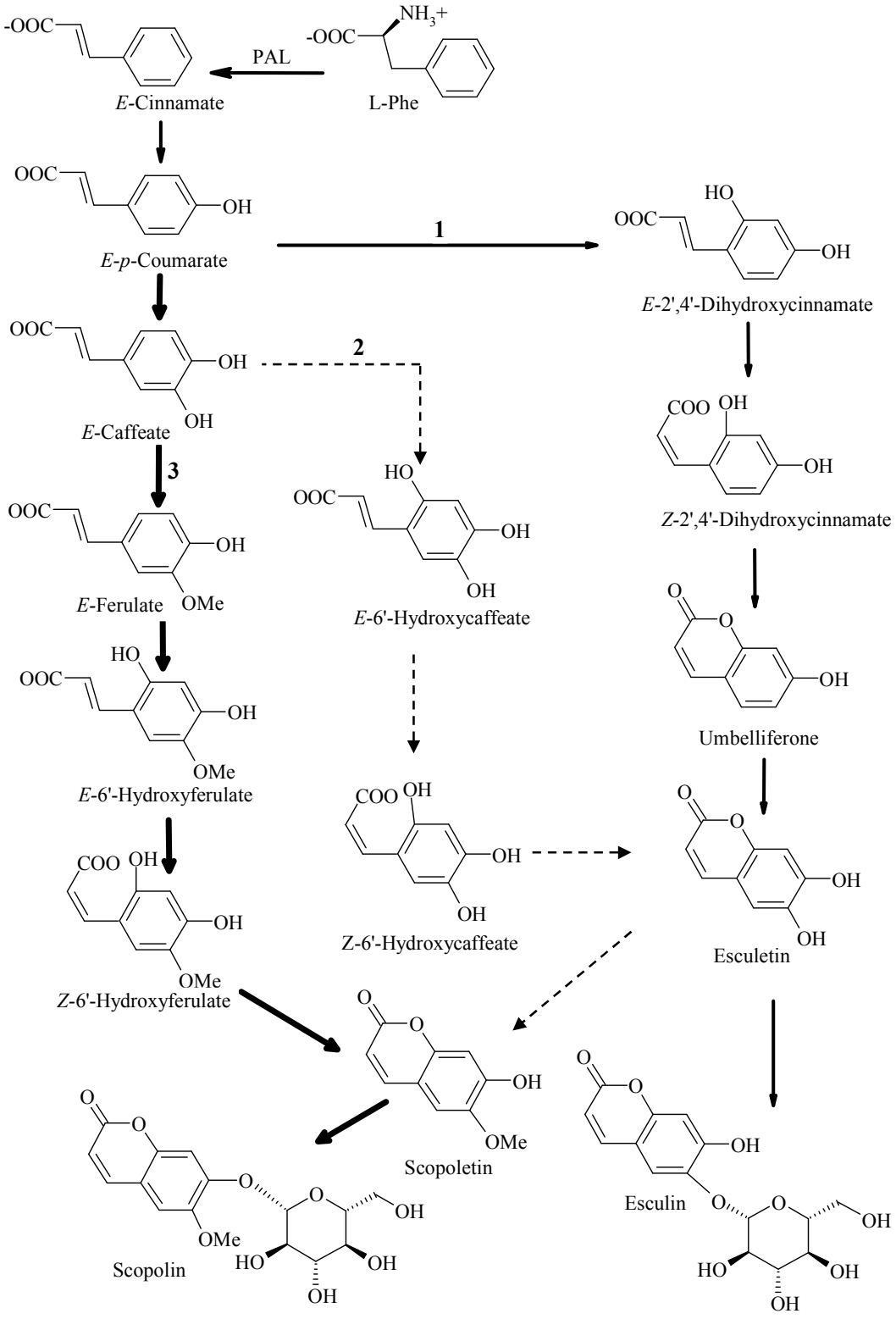
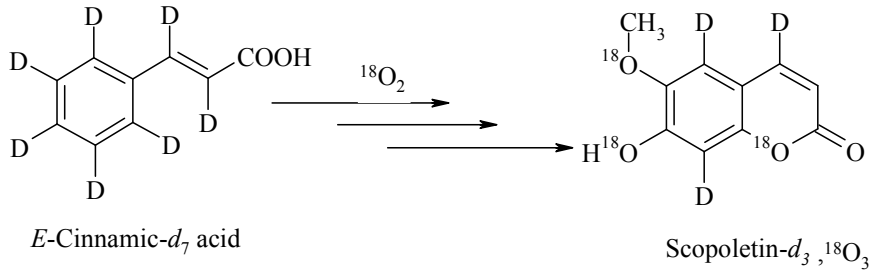


Fig. 6.2. Results of feeding experiments support route 3 as the major pathway and route 1 as a minor pathway for the biosynthesis of scopoletin in cassava roots during PPD.

Competition feeding experiments were carried out by feeding the root samples with deuteriated *E*-cinnamic-(d_5 and d_7) acid alone as a control and mixed with putative unlabelled intermediates along the scopoletin biosynthetic pathway which could compete with deuteriated *E*-cinnamic acid.

Feeding with p-coumaric-2-^{13}C acid, caffeic-2-^{13}C acid and ferulic-2-^{13}C acid showed an increase in the percentage of labelled scopoletin and scopolin above the natural abundance for all three precursors. Ferulic-2-^{13}C acid incorporation confirms that pathway 3 is involved in the biosynthesis of scopoletin in cassava roots during PPD. Feeding with ^{18}O-umbelliferone and ^{18}O-esculetin showed very low incorporation of ^{18}O into scopoletin and scopolin but gave good incorporation into esculin. These results confirm that the major pathway for the biosynthesis of scopoletin and scopolin is through ferulic acid and that biosynthesis of esculetin and esculin is through umbelliferone.

In order to investigate the origin of the lactone ether oxygen in coumarins biosynthesised in cassava roots during PPD, feeding experiments with C^{18}O$_2$-enriched *E*-cinnamic and ferulic acids were carried out. Scopoletin and scopolin biosynthesised were scopoletin-2-^{18}O and scopolin-2-^{18}O, i.e. scopoletin and scopolin only enriched in the carbonyl oxygen. No HR MS peaks corresponding to scopoletin-1,2-^{18}O$_2$ and scopolin-1,2-^{18}O$_2$ were found; this is evidence that their biosynthesis in cassava roots during PPD occurs via *o*-hydroxylation and not via a spirodienone intermediate where both ^{18}O-atoms would have been incorporated in the final product. Separately, we put the harvested roots in an atmosphere rich in ^{18}O$_2$, the scopoletin and scopolin labelling patterns resulting from deuteriated cinnamic-d_7 acid feeding were obtained and compared with those from untreated roots kept in the same vacuum desiccator. The biosynthesis of scopoletin-d_3,^{18}O$_3$ (Fig. 6.3) confirms that their biosynthesis in cassava roots during PPD occurs via *o*-hydroxylation.



E-Cinnamic-d_7 acid

Scopoletin-d_3 ,^{18}O$_3$

Fig. 6.3. Results of feeding experiments in an atmosphere of 20% ^{18}O$_2$, 80%N$_2$

Scopolin accumulates in cassava roots during PPD, and the last step in the biosynthesis of this glycoside is the addition of a glucose unit to scopoletin, catalysed by glucosyltransferases. In order to gain further data on this step, glucosyltransferase clones were isolated from a cassava PPD-related cDNA library and the expression of the corresponding genes was assayed in both fresh and deteriorated cassava roots. Two new glucosyltransferases that expressed in cassava root during PPD were identified; they had previously been identified from cassava cotyledons. Screening the cassava PPD-related cDNA library, using the labelled tobacco scopoletin glucosyltransferase TOGT1 probe, led to the identification of a novel glucosyltransferase. Strong evidence that these genes are expressed in the cassava storage root during PPD came with the isolation of three glucosyltransferases from a PPD-related cDNA library. To confirm this and to compare expression in the root and leaves RT-PCR expression studies were performed. No significant differences are apparent over the time-course of PPD or between roots and leaves and thus these genes may not play an important role specifically during PPD.

These studies were interdisciplinary across the chemical and biological sciences, from different chromatographic techniques to isolate natural products from extracts to detailed NMR and HR MS for their identification. Different stable isotopically labelled (^{2}H, ^{18}O and ^{13}C) intermediates were synthesised and characterised by ^{1}H, ^{2}H, ^{13}C NMR and HR ESI MS, and a general feeding method was developed to study secondary metabolite production. Three glucosyltransferase genes were isolated and identified using bioinformatics methods such as sequence alignment and trees designed to compare sequence relationships. Degenerate primers were designed and used to amplify genes by PCR. A cDNA library was screened using labelled probes. RT PCR was performed to compare gene expression.

It has been established by experiments that the accumulation of coumarins in cassava during PPD makes cassava a good model for studying the biosynthesis of scopoletin. The *E*-*Z*-isomerisation step is enzymatic and not photochemical in cassava roots during PPD. The predominant pathway for scopoletin biosynthesis is through ferulic acid while the pathway through umbelliferone operates for esculetin and esculin biosynthesis. The major pathway in the lactonisation step is through *o*-hydroxylation (not via a spirodienone intermediate). Two new cassava root glucosyltransferases and a novel glucosyltransferase were isolated from a cassava PPD-related cDNA library.

Future work

Secondary metabolites were identified mainly from cassava roots while other organs such as the aerial parts of cassava secondary metabolites need more investigation. The biological activities of these metabolites could be tested for antimicrobial, antioxidant, anti-inflammatory and anticancer activities.

We confirmed that the major pathway for the biosynthesis of scopoletin and scopolin is through ferulic acid and that biosynthesis of esculetin and esculin is through umbelliferone. Enzymes involved in these biosynthesis pathways are phenylalanine ammonia lyase, cinnamate-4'-hydroxylase, *p*-coumaroyl shikimate/quinate 3'-hydroxylase, caffeoyl-CoA *O*-methyltransferase and ferulate *o*-hydroxylase. Genes that code for these enzymes have been identified from different plants but not from cassava. The analysis of their sequences is in Chapter 5. Cinnamate-4'-hydroxylases, *p*-coumaroyl shikimate/quinate 3'-hydroxylases, caffeoyl-CoA *O*-methyltransferases and ferulate *o*-hydroxylases could be identified from a cassava PPD-related cDNA library and their expression could be compared between fresh and deteriorated root. Genes which show dramatic increase in their expression in the deteriorated root compared to the fresh root could be knocked down. The effect of such knocking down could affect the accumulation of hydroxycoumarins after harvesting cassava roots. The knocking down of these enzymes could throw light on the role of hydroxycoumarin accumulation in the overall process of PPD. Also missing enzymes along the biosynthetic pathway, such as isomerase, could be identified. Candidates could be tested by expressing them in *E. coli* or yeast and feeding them with a range of potential substrates.

Investigation of the *E*-*Z*-isomerisation of the C=C double bond in the biosynthesis of scopoletin and scopolin in cassava roots during PPD involves the specific exchange of the hydrogen atom at position 2 of *E*-cinnamic acid and thus is likely to be enzyme catalysed and not photochemical (Chapter 3). Different mechanisms could explain this specific exchange of the hydrogen atom at position 2 of *E*-cinnamic acid as the nucleophilic addition of glutathione as in maleylacetone *Z*-*E*-isomerisation or water addition across the C=C double bond may occur as in the second step of the well known β-oxidation of fatty acids. Since in the literature there is nothing about the mechanism of the *E*-*Z* isomerisation step of cinnamic acid derivatives, this mechanism

could be further investigated in more detail at both an enzymatic and genetic level. Glutathione transferase inhibitors such as (ethacrynic acid and tocopherols) could be used to investigate this step.

During our feeding experiments we used different stable isotopically labelled intermediates and labelled scopoletin and scopolin structures were identified using HR MS which is a highly sensitive technique, but for further analysis more labelled scopoletin and scopolin could be isolated (1-10 mg) and identified by NMR.

During the investigation of *o*-hydroxylation step cassava roots were fed with $C^{18}O_2$-enriched *E*-cinnamic and ferulic acids. Scopoletin and scopolin biosynthesised were scopoletin-2-^{18}O and scopolin-2-^{18}O which were identified by HR MS (Chapter 4). $C^{17}O_2$-enriched *E*-cinnamic and ferulic acids could be synthesised using $H_2{}^{17}O$. Scopoletin and scopolin biosynthesised should be scopoletin-2-^{17}O and scopolin-2-^{17}O which could be identified by (^{17}O observe) NMR in addition to HR MS.

It has been established by experiments that the accumulation of coumarins in cassava during PPD makes cassava a good model for studying the biosynthesis of scopoletin. Similarly cassava could be a good model for studying of the biosynthesis of diterpenes and flavanols.

There were no significant expression differences between the time-course of PPD or between roots and leaves of the identified glucosyltransferases and thus these genes may not play important roled specifically during PPD. Other glucosyltransferases could be identified which may play important roles specifically during PPD.

The interdisciplinary studies across the chemical and biological sciences, from different chromatographic techniques to isolation of natural products from extracts, to detailed NMR and HR MS for their identification could be applied for other economic, nutritious or medicinal plants. Also bioinformatics methods such as sequence alignment and neighbour joining trees could be designed and used to compare genetic sequence relationships. Different genes involved in the biosynthesis of different secondary metabolites will thereby be identified.

References

Abd El-Mawla A.M.A., Beerhues L., 2002. Benzoic acid biosynthesis in cell cultures of *Hypericum androsaemum*. Planta 214, 727-733.

Achnine L., Huhman D.V., Farag M.A., Sumner L.W., Blount J.W., Dixon R.A., 2005. Genomics-based selection and functional characterization of triterpene glycosyltransferases from the model legume *Medicago truncatula*. Plant J. 41, 875-887.

Adams M., Efferth T., Bauer R., 2006. Activity-guided isolation of scopoletin and isoscopoletin, the inhibitory active principles towards CCRF-CEM leukaemia cells and multi-drug resistant CEM/ADR5000 cells, from *Artemisia argyi*. Planta Med. 72, 862-864.

Adewusi S.R.A., Bradbury J.H., 1993. Carotenoids in cassava - comparison of open-column and HPLC methods of analysis. J. Sci. Food. Agric. 62, 375-383.

Alexandrov N.N., Troukhan M.E., Brover V.V., Tatarinova T., Flavell R.B., Feldmann K.A., 2006. Features of *Arabidopsis* genes and genome discovered using full-length cDNAs. Plant Mol. Biol. 60, 69-85.

Anterola A.M., Jeon J.H., Davin L.B., Lewis N.G., 2002. Transcriptional control of monolignol biosynthesis in *Pinus taeda* - Factors affecting monolignol ratios and carbon allocation in phenylpropanoid metabolism. J. Biol. Chem. 277, 18272-18280.

Arcos M.L.B., Cremaschi G., Werner S., Coussio J., Ferraro G., Anesini C., 2006. *Tilia cordata* Mill. extracts and scopoletin (isolated compound): Differential cell growth effects on lymphocytes. Phytother. Res. 20, 34-40.

Averre C.W., 1967. Vascular streaking of stored cassava roots. Int. Symp. Trop. Root Crops 2, 31.

Bahnson B.J., Anderson V.E., Petsko G.A., 2002. Structural mechanism of enoyl-CoA hydratase: Three atoms from a single water are added in either an E1cb stepwise or concerted fashion. Biochemistry 41, 2621-2629.

BellLelong D.A., Cusumano J.C., Meyer K., Chapple C., 1997. Cinnamate-4'-hydroxylase expression in *Arabidopsis* - Regulation in response to development and the environment. Plant Physiol. 113, 729-738.

Bernards M.A., Fleming W.D., Llewellyn D.B., Priefer R., Yang X.L., Sabatino A., Plourde G.L., 1999. Biochemical characterization of the suberization-associated anionic peroxidase of potato. Plant Physiol. 121, 135-145.

Betz C., McCollum T.G., Mayer R.T., 2001. Differential expression of two cinnamate 4-hydroxylase genes in 'Valencia' orange (*Citrus sinensis* Osbeck). Plant Mol. Biol. 46, 741-748.

Bouis H.E., 2003. Micronutrient fortification of plants through plant breeding: can it improve nutrition in man at low cost? Proc. Nutr. Soc. 62, 403-411.

Bourgaud F., Hehn A., Larbat R., Doerper S., Gontier E., Kellner S., Matern U., 2006. Biosynthesis of coumarins in plants: a major pathway still to be unravelled for cytochrome P450 enzymes. Phytochemistry Rev. 5, 293-308.

Bowles D., Isayenkova J., Lim E.K., Poppenberger B., 2005. Glycosyltransferases: managers of small molecules. Curr. Opin. Plant Biol. 8, 254-263.

Bowles D., Lim E.K., Poppenberger B., Vaistij F.E., 2006. Glycosyltransferases of lipophilic small molecules. Annu. Rev. Plant Biol. 57, 567-597.

Bradbury J.H., 2006. Simple wetting method to reduce cyanogen content of cassava flour. J. Food Compost. Anal. 19, 388-393.

Brown D., Owens R.A., McMahon V.A., 1975. Phenolic constituents of *Artemesia tridentata* spp. *vaseyana*. Phytochemistry 14, 1083-1084.

Brown G.R., Gill G.P., Kuntz R.J., Langley C.H., Neale D.B., 2004. Nucleotide diversity and linkage disequilibrium in loblolly pine. Proc. Natl. Acad. Sci. U. S. A. 101, 15255-15260.

Brown SA, 1981. Coumarins. In E.E. Conn, ed, The biochemistry of plants. A compehensive Treatise. Academic press, London, pp 269-300.

Brown S.A., 1985. Biosynthesis of 6,7-dihydroxycoumarin in *Cichorium intybus*. Can. J. Biochem. Cell Biol. 63, 292-295.

Brown S.A., March R.E., Rivett D.E.A., Thompson H.J., 1988. Intermediates in the formation of puberulin by *Agathosma puberula*. Phytochemistry 27, 391-395.

Brown S.A., Rivett D.E.A., Thompson H.J., 1984. Elaboration of the 6,7,8 oxygenation pattern in simple coumarins - biosynthesis of puberulin in *Agathosma Puberula* Fourc. Z. Naturforsch. ,C: Biosci. 39, 31-37.

Bruzik K.S., Nyholm P.G., 1997. NMR study of the conformation of galactocerebroside in bilayers and solution: Galactose reorientation during the metastable- stable gel transition. Biochemistry 36, 566-575.

Bull S.E., Briddon R.W., Sserubombwe W.S., Ngugi K., Markham P.G., Stanley J., 2006. Genetic diversity and phylogeography of cassava mosaic viruses in Kenya. J. Gen. Virol. 87, 3053-3065.

Bunton C.A., Kenner G.W., Robinson M.J., Webster B.R., 1963. Experiments related to biosynthesis of novobiocin and other coumarins. Tetrahedron 19, 1001-1010.

Busam G., Junghanns K.T., Kneusel R.E., Kassemeyer H.H., Matern U., 1997. Characterization and expression of caffeoyl-coenzyme A 3-*O*-methyltransferase proposed for the induced resistance response of *Vitis vinifera* L. Plant Physiol. 115, 1039-1048.

Buschmann H., Reilly K., Rodriguez M.X., Tohme J., Beeching J.R., 2000a. Hydrogen peroxide and flavan-3-ols in storage roots of cassava (*Manihot esculenta* crantz) during postharvest deterioration. J. Agric. Food Chem. 48, 5522-5529.

Buschmann H., Rodriguez M.X., Tohme J., Beeching J.R., 2000b. Accumulation of hydroxycoumarins during post-harvest deterioration of tuberous roots of Cassava (*Manihot esculenta* Crantz). Ann. Bot. 86, 1153-1160.

Cardoso A.P., Mirione E., Ernesto M., Massaza F., Cliff J., Haque M.R., Bradbury J.H., 2005. Processing of cassava roots to remove cyanogens. J. Food Compost. Anal. 18, 451-460.

Catinot J., Buchala A., Abou-Mansour E., Metraux J.P., 2008. Salicylic acid production in response to biotic and abiotic stress depends on isochorismate in *Nicotiana benthamiana*. FEBS Lett. 582, 473-478.

Ceballos H., Sanchez T., Chavez A.L., Iglesias C., Debouck D., Mafla G., Tohme J., 2006. Variation in crude protein content in cassava (*Manihot esculenta* Crantz) roots. J. Food Compost. Anal. 19, 589-593.

Cespedes C.L., Avila J.G., Martinez A., Serrato B., Calderon-Mugica J.C., Salgado-Garciglia R., 2006. Antifungal and antibacterial activities of *Mexican tarragon* (*Tagetes lucida*). J. Agric. Food Chem. 54, 3521-3527.

Chandrasekaran S., Wilson W.D., Boykin D.W., 1984. ^{17}O NMR studies on polycyclic quinones, hydroxyquinones and related cyclic ketones: Models for anthracycline intercalators. Org. Magn. Reson. 22, 757-760.

Chang W.S., Chiang H.C., 1995. Structure-activity relationship of coumarins in xanthine oxidase inhibition. Anticancer Res. 15, 1969-1973.

Charles A.L., Sriroth K., Huang T.C., 2005. Proximate composition, mineral contents, hydrogen cyanide and phytic acid of 5 cassava genotypes. Food Chem. 92, 615-620.

Chellappan P., Masona M.V., Vanitharani R., Taylor N.J., Fauquet C.M., 2004. Broad spectrum resistance to ssDNA viruses associated with transgene-induced gene silencing in cassava. Plant Mol. Biol. 56, 601-611.

Chen A.H., Chai Y.R., Li J.N., Chen L., 2007. Molecular cloning of two genes encoding cinnamate 4-hydroxylase (C4H) from oilseed rape (*Brassica napus*). J. Biochem. Mol. Biol. 40, 247-260.

Chong J., Baltz R., Schmitt C., Beffa R., Fritig B., Saindrenan P., 2002. Downregulation of a pathogen-responsive tobacco UDP-Glc : phenylpropanoid glucosyltransferase reduces scopoletin glucoside accumulation, enhances oxidative stress, and weakens virus resistance. Plant Cell 14, 1093-1107.

Clapp R.C., Bissett F.H., Coburn R.A., Long L., 1966. Cyanogenesis in Manioc: linamarin and isolinamarin. Phytochemistry 5, 1323-1326.

Clayden J, Greeves N, Warren S, Wothers P, 2001. Organic chemistry. Oxford Universiy Press, Oxford, pp 702-703.

Cochrane F.C., Davin L.B., Lewis N.G., 2004. The *Arabidopsis* phenylalanine ammonia lyase gene family: kinetic characterization of the four PAL isoforms. Phytochemistry 65, 1557-1564.

Cramer C.L., Edwards K., Dron M., Liang X.W., Dildine S.L., Bolwell G.P., Dixon R.A., Lamb C.J., Schuch W., 1989. Phenylalanine ammonia-lyase gene organization and structure. Plant Mol. Biol. 12, 367-383.

Devereux J., Haeberli P., Smithies O., 1984. A comprehensive set of sequence-analysis programs for the vax. Nucleic Acids Res. 12, 387-395.

Dixon D.P., Cole D.J., Edwards R., 2000. Characterisation of a zeta class glutathione transferase from *Arabidopsis thaliana* with a putative role in tyrosine catabolism. Arch. Biochem. Biophys. 384, 407-412.

Do C.T., Pollet B., Thevenin J., Sibout R., Denoue D., Barriere Y., Lapierre C., Jouanin L., 2007. Both caffeoyl Coenzyme A 3-*O*-methyltransferase 1 and caffeic acid *O*-methyltransferase 1 are involved in redundant functions for lignin, flavonoids and sinapoyl malate biosynthesis in *Arabidopsis*. Planta 226, 1117-1129.

Du L.C., Bokanga M., Moller B.L., Halkier B.A., 1995. The biosynthesis of cyanogenic glucosides in roots of cassava. Phytochemistry 39, 323-326.

Dugave C., Demange L., 2003. Cis-trans isomerization of organic molecules and biomolecules: Implications and applications. Chem. Rev. 103, 2475-2532.

Edwards K.G., Stoker J.R., 1967. Biosynthesis of coumarin: The isomerization stage. Phytochemistry 6, 655-661.

Edwards K.G., Stoker J.R., 1968. Biosynthesis of herniarin: The isomerization stage. Phytochemistry 7, 73-77.

Ehlting J., Hamberger B., Million-Rousseau R., Werck-Reichhart D., 2006. Cytochromes P450 in phenolic metabolism. Phytochemistry Rev. 5, 239-270.

Ellis B.E., Amrhein N., 1971. NIH-shift during aromatic orthohydroxylation in higher plants. Phytochemistry 10, 3069-3072.

Farah M.H., Samuelsson G., 1992. Pharmacologically active phenylpropanoids from *Senra incana*. Planta Med. 58, 14-18.

Fliniaux M.A., GilletManceau F., Marty D., Macek T., Monti J.P., JacquinDubreuil A., 1997. Evaluation of the relation between the endogenous scopoletin and scopolin level of some solanaceous and papaver cell suspensions and their ability to bioconvert scopoletin to scopolin. Plant Sci. 123, 205-210.

Fraissinet-Tachet L., Baltz R., Chong J., Kauffmann S., Fritig B., Saindrenan P., 1998. Two tobacco genes induced by infection, elicitor and salicylic acid encode glucosyltransferases acting on phenylpropanoids and benzoic acid derivatives, including salicylic acid. FEBS Lett. 437, 319-323.

Fregene M., Matsumura H., Akano A., Dixon A., Terauchi R., 2004. Serial analysis of gene expression (SAGE) of host-plant resistance to the cassava mosaic disease (CMD). Plant Mol. Biol. 56, 563-571.

Frick S., Kutchan T.M., 1999. Molecular cloning and functional expression of *O*-methyltransferases common to isoquinoline alkaloid and phenylpropanoid biosynthesis. Plant J. 17, 329-339.

Fritig B., Hirth L., Ourisson G., 1970. Biosynthesis of the coumarins: Scopoletin formation in tobacco tissue cultures. Phytochemistry 9, 1963-1975.

Gachon C., Baltz R., Saindrenan P., 2004. Over-expression of a scopoletin glucosyltransferase in *Nicotiana tabacum* leads to precocious lesion formation during the hypersensitive response to tobacco mosaic virus but does not affect virus resistance. Plant Mol. Biol. 54, 137-146.

Galabov A.S., Iosifova T., Vassileva E., Kostova I., 1996. Antiviral activity of some hydroxycoumarin derivatives. Z. Naturforsch., C, J. Biosci. 51, 558-562.

Gang D.R., Beuerle T., Ullmann P., Werck-Reichhart D., Pichersky E., 2002. Differential production of meta hydroxylated phenylpropanoids in sweet basil peltate glandular trichomes and leaves is controlled by the activities of specific acyltransferases and hydroxylases. Plant Physiol. 130, 1536-1544.

Garcia D., Sanier C., Macheix J.J., Dauzac J., 1995. Accumulation of scopoletin in *Hevea brasiliensis* infected by *Microcyclus ulei* (Henn,P.) V Arx and evaluation of its fungitoxicity for 3 leaf pathogens of rubber trees. Physiol. Mol. Plant Pathol. 47, 213-223.

Gasson M.J., Kitamura Y., McLauchlan W.R., Narbad A., Parr A.J., Lindsay E., Parsons H., Payne J., Rhodes M.J.C., Walton N.J., 1998. Metabolism of ferulic acid to vanillin - A bacterial gene of the enoyl-SCoA hydratase/isomerase superfamily encodes an enzyme for the hydration and cleavage of a hydroxycinnamic acid SCoA thioester. J. Biol. Chem. 273, 4163-4170.

Giesemann A., Biehl B., Lieberei R., 1986. Identification of scopoletin as a phytoalexin of the rubber tree *Hevea brasiliensis*. J. Phytopathol. 117, 373-376.

Gilani A.H., Janbaz K.H., Shah B.H., 1998. Esculetin prevents liver damage induced by paracetamol and CCl4. Pharmacol. Res. 37, 31-35.

Gómez-Vásquez R., Day R., Buschmann H., Randles S., Beeching J.R., Cooper R.M., 2004. Phenylpropanoids, phenylalanine ammonia lyase and peroxidases in elicitor-challenged cassava (*Manihot esculenta*) suspension cells and leaves. Ann. Bot. 94, 87-97.

Goy P.A., Signer H., Reist R., Aichholz R., Blum W., Schmidt E., Kessmann H., 1993. Accumulation of scopoletin is associated with the high disease resistance of the hybrid *Nicotiana-glutinosa* X *Nicotiana-debneyi*. Planta 191, 200-206.

Grisebach H., Ollis W.D., 1961. Biogenetic relationships between coumarins, flavonoids, isoflavonoids, and rotenoids. Experientia 17, 4.

Guantai A.N., Addae-Mensah I., 1999. Cardiovascular effect of Artemisia afra and its constituents. Pharm. Biol. 37, 351-356.

Guillet-Claude C., Birolleau-Touchard C., Manicacci D., Fourmann M., Barraud S., Carret V., Martinant J.P., Barriere Y., 2004. Genetic diversity associated with variation in silage corn digestibility for three *O*-methyltransferase genes involved in lignin biosynthesis. Theor. Appl. Genet. 110, 126-135.

Gutierrez M.C., Parry A., Tena M., Jorrin J., Edwards R., 1995. Abiotic elicitation of coumarin phytoalexins in sunflower. Phytochemistry 38, 1185-1191.

Hardick D.J., Blagbrough I.S., Potter B.V.L., 1996. Isotopic enrichment by asymmetric deuteriation. An investigation of the synthesis of deuteriated (S)-(-)-methylsuccinic acids from itaconic acid. J. Am. Chem. Soc. 118, 5897-5903.

Hertweck C., Jarvis A.P., Xiang L.K., Moore B.S., Oldham N.J., 2001. A mechanism of benzoic acid biosynthesis in plants and bacteria that mirrors fatty acid β-oxidation. Chembiochem 2, 784-786.

Hirata T., Shimoda K., Fujino T., Ohta S., 2000. Biotransformation of hydroxycoumarins by the cultured cells of *Nicotiana tabacum*. J. Mol. Catal. , B Enzym. 10, 477-481.

Hirotani M., Kuroda R., Suzuki H., Yoshikawa T., 2000. Cloning and expression of UDP-glucose: flavonoid 7-O-glucosyltransferase from hairy root cultures of *Scutellaria baicalensis*. Planta 210, 1006-1013.

Horvath D.M., Chua N.H., 1996. Identification of an immediate-early salicylic acid-inducible tobacco gene and characterization of induction by other compounds. Plant Mol. Biol. 31, 1061-1072.

Huang H.C., Lai M.W., Wang H.R., Chung Y.L., Hsieh L.M., Chen C.C., 1993. Antiproliferative effect of esculetin on vascular smooth-muscle cells - possible roles of signal-transduction pathways. Eur. J. Pharmacol. 237, 39-44.

Hughes J., Hughes M.A., 1994. Multiple secondary plant product UDP-glucose glucosyltransferase genes expressed in cassava (*Manihot esculenta* Crantz) cotyledons. DNA Seq. 5, 41-49.

Hundle B.S., Obrien D.A., Alberti M., Beyer P., Hearst J.E., 1992. Functional expression of zeaxanthin glucosyltransferase from *Erwinia herbicola* and a proposed uridine-diphosphate binding-site. Proc. Natl. Acad. Sci. U. S. A. 89, 9321-9325.

Ibdah M., Zhang X.H., Schmidt J., Vogt T., 2003. A novel Mg2+-dependent *O*-methyltransferase in the phenylpropanoid metabolism of *Mesembryanthemum crystallinum*. J. Biol. Chem. 278, 43961-43972.

Inoue K., Sewalt V.J.H., Ballance G.M., Ni W.T., Sturzer C., Dixon R.A., 1998. Developmental expression and substrate specificities of alfalfa caffeic acid 3-*O*-methyltransferase and caffeoyl coenzyme A 3-*O*-methyltransferase in relation to lignification. Plant Physiol. 117, 761-770.

Janwitayanuchit W., Suwanborirux K., Patarapanich C., Pummangura S., Lipipun V., Vilaivan T., 2003. Synthesis and anti-herpes simplex viral activity of monoglycosyl diglycerides. Phytochemistry 64, 1253-1264.

Ji R., Chen Z.X., Corvini P.F.X., Kappler A., Brune A., Haider K., Schaffer A., 2005. Synthesis of [^{13}C]- and [^{14}C]-labeled phenolic humus and lignin monomers. Chemosphere 60, 1169-1181.

Jin S.J., Hoppel C.L., Tserng K.Y., 1992. Incomplete fatty acid oxidation. The production and epimerization of 3-hydroxy fatty acids. J. Biol. Chem. 267, 119-125.

Jones P., Vogt T., 2001. Glycosyltransferases in secondary plant metabolism: tranquilizers and stimulant controllers. Planta 213, 164-174.

Jones P.R., Moller B.L., Hoj P.B., 1999. The UDP-glucose : p-hydroxymandelonitrile-*O*-glucosyltransferase that catalyzes the last step in synthesis of the cyanogenic glucoside dhurrin in *Sorghum bicolor* - Isolation, cloning, heterologous expression, and substrate specificity. J. Biol. Chem. 274, 35483-35491.

Jorgensen K., Bak S., Busk P.K., Sorensen C., Olsen C.E., Puonti-Kaerlas J., Moller B.L., 2005. Cassava plants with a depleted cyanogenic glucoside content in leaves and tubers. Distribution of cyanogenic glucosides, their site of synthesis and transport, and blockage of the biosynthesis by RNA interference technology. Plant Physiol. 139, 363-374.

Kai K., Mizutani M., Kawamura N., Yamamoto R., Tamai M., Yamaguchi H., Sakata K., Shimizu B.I., 2008. Scopoletin is biosynthesized via *ortho*-hydroxylation of feruloyl-

CoA by an 2-oxoglutarate dependent dioxygenase in *Arabidopsis thaliana*. Plant J. 55, 989-999.

Kai K., Shimizu B., Mizutani M., Watanabe K., Sakata K., 2006. Accumulation of coumarins in *Arabidopsis thaliana*. Phytochemistry 67, 379-386.

Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S., Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T., Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A., Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M., Yasuda M., Tabata S., 2001. Complete genomic sequence of the filamentous nitrogen-fixing *Cyanobacterium anabaena* sp strain PCC 7120. DNA Res. 8, 205-213.

Kaneko T., Tahara S., Takabayashi F., 2007. Inhibitory effect of natural coumarin compounds, esculetin and esculin, on oxidative DNA damage and formation of aberrant crypt foci and tumors induced by 1,2-dimethylhydrazine in rat colons. Biol. Pharm. Bull. 30, 2052-2057.

Kang T.H., Pae H.O., Jeong S.J., Yoo J.C., Choi B.M., Jun C.D., Chung H.T., Miyamoto T., Higuchi R., Kim Y.C., 1999. Scopoletin: An inducible nitric oxide synthesis inhibitory active constituent from *Artemisia feddei*. Planta Med. 65, 400-403.

Kapitonov D., Yu R.K., 1999. Conserved domains of glycosyltransferases. Glycobiology 9, 961-978.

Kawaii S., Tomono Y., Katase E., Ogawa K., Yano M., 2000. Effect of coumarins on HL-60 cell differentiation. Anticancer Res. 20, 2505-2512.

Kawaii S., Tomono Y., Ogawa K., Sugiura M., Yano M., Yoshizawa Y., 2001. The antiproliferative effect of coumarins on several cancer cell lines. Anticancer Res. 21, 917-923.

Kawase M., Sakagami H., Hashimoto K., Tani S., Hauer H., Chatterjee S.S., 2003. Structure-cytotoxic activity relationships of simple hydroxylated coumarins. Anticancer Res. 23, 3243-3246.

Kemp B.P., Beeching J.R., Cooper R.M., 2005. cDNA-AFLP reveals genes differentially expressed during the hypersensitive response of cassava. Mol. Plant Pathol. 6, 113-123.

Kemp B.P., Horne J., Bryant A., Cooper R.M., 2004. *Xanthomonas axonopodis* pv. manihotis gumD gene is essential, for EPS production and pathogenicity and enhances epiphytic survival on cassava (*Manihot esculenta*). Physiol. Mol. Plant Pathol. 64, 209-218.

Khan M.S.Y., Javed K., Khan M.H., Shamsi M.A., Siddiqui A.A., 1991. α-Amyrin derivatives from *Corchorus depressus*. Phytochemistry 30, 1989-1992.

Kim A.R., Zou Y.N., Park T.H., Shim K.H., Kim M.S., Kim N.D., Kim J.D., Bae S.J., Choi J.S., Chung H.Y., 2004. Active components from *Artemisia iwayomogi* displaying $ONOO^-$ scavenging activity. Phytother. Res. 18, 1-7.

Kim B.G., Lee Y., Hur H.G., Lim Y., Ahn J.H., 2006a. Production of three *O*-methhylated esculetins with *Escherichia coli* expressing *O*-methyltransferase from poplar. Biosci. Biotechnol. Biochem. 70, 1269-1272.

Kim E.K., Kwon K.B., Shin B.C., Seo E.A., Lee Y.R., Kim J.S., Park J.W., Park B.H., Ryu D.G., 2005. Scopoletin induces apoptosis in human promyeloleukemic cells, accompanied by activations of nuclear factor kappa B and caspase-3. Life Sci. 77, 824-836.

Kim J.S., Kim J.C., Shim S.H., Lee E.J., Jin W., Bae K., Son K.H., Kim H.P., Kang S.S., Chang H.W., 2006b. Chemical constituents of the root of *Dystaenia takeshimana* and their anti-inflammatory activity. Arch. Pharm. Res. 29, 617-623.

Kim N.M., Kim J., Chung H.Y., Choi J.S., 2000. Isolation of luteolin 7-O-rutinoside and esculetin with potential antioxidant activity from the aerial parts of *Artemisia montana*. Arch. Pharm. Res. 23, 237-239.

King N.L.R., Bradbury J.H., 1995. Bitterness of cassava - identification of a new apiosyl glucoside and other compounds that affect its bitter taste. J. Sci. Food. Agric. 68, 223-230.

Klaus D., Hartel H., Fitzpatrick L.M., Froehlich J.E., Hubert J., Benning C., Dormann P., 2002. Digalactosyldiacylglycerol synthesis in chloroplasts of the *Arabidopsis dgd*1 mutant. Plant Physiol. 128, 885-895.

Kobayashi S., Ishimaru M., Ding C.K., Yakushiji H., Goto N., 2001. Comparison of UDP-glucose : flavonoid 3-O-glucosyltransferase (UFGT) gene sequences between white grapes (*Vitis vinifera*) and their sports with red skin. Plant Sci. 160, 543-550.

Koenigs P.M., Faust B.C., Porter N.A., 1993. Photochemistry of enzyme-bound cinnamoyl derivatives. J. Am. Chem. Soc. 115, 9371-9379.

Kohara A., Nakajima C., Hashimoto K., Ikenaga T., Tanaka H., Shoyama Y., Yoshida S., Muranaka T., 2005. A novel glucosyltransferase involved in steroid saponin biosynthesis in *Solanum aculeatissimum*. Plant Mol. Biol. 57, 225-239.

Kojima H., Sato N., Hatano A., Ogura H., 1990. Constituents of the Labiatae plants. 5. Sterol glucosides from *Prunella vulgaris*. Phytochemistry 29, 2351-2355.

Kojima M., Kimura H., Ohnishi M., Fujino Y., Ito S., 1991. Distribution and characterization of diglycosyldiacylglycerol isomers with different anomeric configuration in higher plants. Phytochemistry 30, 1165-1168.

Kollenz G., Sterk H., Hutter G., 1991. Mechanistic investigations aided by isotopic labeling.10. Investigations of novel furan-2,3-dione rearrangements by oxygen-17 labeling. J. Org. Chem. 56, 235-239.

Koukol J., Conn E.E., 1961. Metabolism of aromatic compounds in higher plants. IV. purification and properties of phenylalanine deaminase of *Hordeum vulgare*. J. Biol. Chem. 236, 2692-2698.

Kozminski W., Nanz D., 1996. Sensitive measurement and unambiguous assignment of long-range ^{13}C, ^{13}C coupling constants at natural isotope abundance. Journal of Magnetic Resonance Series A 122, 245-247.

Kubo A., Arai Y., Nagashima S., Yoshikawa T., 2004. Alteration of sugar donor specificities of plant glycosyltransferases by a single point mutation. Arch. Biochem. Biophys. 429, 198-203.

Kuo H.C., Lee H.J., Hu C.C., Shun H.I., Tseng T.H., 2006. Enhancement of esculetin on Taxol-induced apoptosis in human hepatoma HepG2 cells. Toxicol. Appl. Pharmacol. 210, 55-62.

Lacy A., O'Kennedy R., 2004. Studies on coumarins and coumarin-related compounds to determine their therapeutic role in the treatment of cancer. Curr. Pharm. Des. 10, 3797-3811.

Lee B.C., Lee S.Y., Lee H.J., Sim G.S., Kim J.H., Kim J.H., Cho Y.H., Lee D.H., Pyo H.B., Choe T.B., Moon D.C., Yun Y.P., Hong J.T., 2007. Anti-oxidative and photo-protective effects of coumarins isolated from *Fraxinus chinensis*. Arch. Pharm. Res. 30, 1293-1301.

Lee H., Raskin I., 1999. Purification, cloning, and expression of a pathogen inducible UDP-glucose: Salicylic acid glucosyltransferase from tobacco. J. Biol. Chem. 274, 36637-36642.

Lee K.T., Lillard D.A., 1997. Effects of esculetin as a lipoxygenase inhibitor in soybean extracts. J. Food Lipids 4, 119-127.

Legg J.P., Fauquet C.M., 2004. Cassava mosaic geminiviruses in Africa. Plant Mol. Biol. 56, 585-599.

Leon J., Shulaev V., Yalpani N., Lawton M.A., Raskin I., 1995. Benzoic acid 2-hydroxylase, a soluble oxygenase from tobacco, catalyzes salicylic acid biosynthesis. Proc. Natl. Acad. Sci. U. S. A. 92, 10413-10417.

Lepelley M., Cheminade G., Tremillon N., Simkin A., Caillet V., McCarthy J., 2007. Chlorogenic acid synthesis in coffee: An analysis of CGA content and real-time RT-PCR expression of HCT, HQT, C3H1, and CCoAOMT1 genes during grain development in *C. canephora*. Plant Sci. 172, 978-996.

Li L.G., Osakabe Y., Joshi C.P., Chiang V.L., 1999. Secondary xylem-specific expression of caffeoyl-coenzyme A 3-*O*-methyltransferase plays an important role in the methylation pathway associated with lignin biosynthesis in loblolly pine. Plant Mol. Biol. 40, 555-565.

Li Y., Baldauf S., Lim E.K., Bowles D.J., 2001. Phylogenetic analysis of the UDP-glycosyltransferase multigene family of *Arabidopsis thaliana*. J. Biol. Chem. 276, 4338-4343.

Lim E.K., Baldauf S., Li Y., Elias L., Worrall D., Spencer S.P., Jackson R.G., Taguchi G., Ross J., Bowles D.J., 2003. Evolution of substrate recognition across a multigene family of glycosyltransferases in Arabidopsis. Glycobiology 13, 139-145.

Lim E.K., Bowles D.J., 2004. A class of plant glycosyltransferases involved in cellular homeostasis. EMBO J. 23, 2915-2922.

Lin H.C., Tsai S.H., Chen C.S., Chang Y.C., Lee C.M., Lai Z.Y., Lin C.M., 2008. Structure-activity relationship of coumarin derivatives on xanthine oxidase-inhibiting and free radical-scavenging activities. Biochem. Pharmacol. 75, 1416-1425.

Logemann E., Parniske M., Hahlbrock K., 1995. Modes of expression and common structural features of the complete phenylalanine ammonia-lyase gene family in parsley. Proc. Natl. Acad. Sci. U. S. A. 92, 5905-5909.

Lopez C., Soto M., Restrepo S., Piegu B., Cooke R., Delseny M., Tohme J., Verdier V., 2005. Gene expression profile in response to *Xanthomonas axonopodis* pv. manihotis infection in cassava using a cDNA microarray. Plant Mol. Biol. 57, 393-410.

Lu S.F., Zhou Y.H., Li L.G., Chiang V.L., 2006. Distinct roles of cinnamate 4-hydroxylase genes in Populus. Plant Cell Physiol. 47, 905-914.

Lunkenbein S., Bellido M., Aharoni A., Salentijn E.M.J., Kaldenhoff R., Coiner H.A., Munoz-Blanco J., Schwab W., 2006. Cinnamate metabolism in ripening fruit. Characterization of a UDP-glucose: Cinnamate glucosyltransferase from strawberry. Plant Physiol. 140, 1047-1058.

Ma G.B., Hayes S.E., 2004. Microwave-enhanced, solvent-free synthesis of singly and doubly ^{13}C-labelled *trans*-cinnamic acid at the α- and β-carbon positions. J. Labelled Compd. Radiopharm. 47, 895-901.

MacDonald M.J., D'Cunha G.B., 2007. A modern view of phenylalanine ammonia lyase. Biochem. Cell Biol 85, 273-282.

Mackenzie P.I., Owens I.S., Burchell B., Bock K.W., Bairoch A., Belanger A., FournelGigleux S., Green M., Hum D.W., Iyanagi T., Lancet D., Louisot P., Magdalou J., Chowdhury J.R., Ritter J.K., Schachter H., Tephly T.R., Tipton K.F., Nebert D.W., 1997. The UDP glycosyltransferase gene superfamily: Recommended nomenclature update based on evolutionary divergence. Pharmacogenetics 7, 255-269.

Maes D., Vervisch S., Debenedetti S., Davio C., Mangelinckx S., Giubellina N., De Kimpe N., 2005. Synthesis and structural revision of naturally occurring ayapin derivatives. Tetrahedron 61, 2505-2511.

Mahesh V., Million-Rousseau R., Ullmann P., Chabrillange N., Bustamante J., Mondolot L., Morant M., Noirot M., Hamon S., de Kochko A., Werck-Reichhart D., Campa C., 2007. Functional characterization of two *p*-coumaroyl ester 3 '-hydroxylase genes from coffee tree: evidence of a candidate for chlorogenic acid biosynthesis. Plant Mol. Biol. 64, 145-159.

Manitto P., Monti D., Gramatic P., 1973. Preparation of cinnamic acids labeled with deuterium or tritium at α-Position. J. Chem. Soc. Chem. Commun. 563-564.

Manuele M.G., Ferraro G., Arcos M.L.B., Lopez P., Cremaschi G., Anesini C., 2006. Comparative immunomodulatory effect of scopoletin on tumoral and normal lymphocytes. Life Sci. 79, 2043-2048.

Marita J.M., Ralph J., Hatfield R.D., Guo D.J., Chen F., Dixon R.A., 2003. Structural and compositional modifications in lignin of transgenic alfalfa down-regulated in caffeic acid 3-*O*-methyltransferase and caffeoyl coenzyme A 3-*O*-methyltransferase. Phytochemistry 62, 53-65.

Martz F., Maury S., Pincon G., Legrand M., 1998. cDNA cloning, substrate specificity and expression study of tobacco caffeoyl-CoA 3-*O*-methyltransferase, a lignin biosynthetic enzyme. Plant Mol. Biol. 36, 427-437.

Masamoto Y., Ando H., Murata Y., Shimoishi Y., Tada M., Takahata K., 2003. Mushroom tyrosinase inhibitory activity of esculetin isolated from seeds of *Euphorbia lathyris* L. Biosci. Biotechnol. Biochem. 67, 631-634.

Masamoto Y., Murata Y., Baba K., Shimoishi Y., Tada M., Takahata K., 2004. Inhibitory effects of esculetin on melanin biosynthesis. Biol. Pharm. Bull. 27, 422-425.

Matern U., 1991. Coumarins and other phenylpropanoid compounds in the defense response of plant cells. Planta Med. 57, S15-S20.

Mativandlela S.P.N., Meyer J.J.M., Hussein A.A., Lall N., 2007. Antitubercular activity of compounds isolated from *Pelargonium sidoides*. Pharm. Biol. 45, 645-650.

Matros A., Mock H.P., 2004. Ectopic expression of a UDP-glucose : phenylpropanoid glucosyltransferase leads to increased resistance of transgenic tobacco plants against infection with potato virus Y. Plant Cell Physiol. 45, 1185-1193.

Matsumoto T., Wu J.Z., Kanamori H., Katayose Y., Fujisawa M., Namiki N., Mizuno H., Yamamoto K., Antonio B.A., Baba T., Sakata K., Nagamura Y., Aoki H., Arikawa K., Arita K., Bito T., Chiden Y., Fujitsuka N., Fukunaka R., Hamada M., Harada C., Hayashi A., Hijishita S., Honda M., Hosokawa S., Ichikawa Y., Idonuma A., Iijima M., Ikeda M., Ikeno M., Ito K., Ito S., Ito T., Ito Y., Ito Y., Iwabuchi A., Kamiya K.,

Karasawa W., Kurita K., Katagiri S., Kikuta A., Kobayashi H., Kobayashi N., Machita K., Maehara T., Masukawa M., Mizubayashi T., Mukai Y., Nagasaki H., Nagata Y., Naito S., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M., Meguro A., Negishi M., Ohta I., Ohta T., Okamoto M., Ono N., Saji S., Sakaguchi M., Sakai K., Shibata M., Shimokawa T., Song J.Y., Takazaki Y., Terasawa K., Tsugane M., Tsuji K., Ueda S., Waki K., Yamagata H., Yamamoto M., Yamamoto S., Yamane H., Yoshiki S., Yoshihara R., Yukawa K., Zhong H.S., Yano M., Sasaki T., Yuan Q.P., Shu O.T., Liu J., Jones K.M., Gansberger K., Moffat K., Hill J., Bera J., Fadrosh D., Jin S.H., Johri S., Kim M., Overton L., Reardon M., Tsitrin T., Vuong H., Weaver B., Ciecko A., Tallon L., Jackson J., Pai G., Van Aken S., Utterback T., Reidmuller S., Feldblyum T., Hsiao J., Zismann V., Iobst S., de Vazeille A.R., Buell C.R., Ying K., Li Y., Lu T.T., Huang Y.C., Zhao Q., Feng Q., Zhang L., Zhu J.J., Weng Q.J., Mu J., Lu Y.Q., Fan D.L., Liu Y.L., Guan J.P., Zhang Y.J., Yu S.L., Liu X.H., Zhang Y., Hong G.F., Han B., Choisne N., Demange N., Orjeda G., Samain S., Cattolico L., Pelletier E., Couloux A., Segurens B., Wincker P., D'Hont A., Scarpelli C., Weissenbach J., Salanoubat M., Quetier F., Yu Y., Kim H.R., Rambo T., Currie J., Collura K., Luo M.Z., Yang T.J., Ammiraju J.S.S., Engler F., Soderlund C., Wing R.A., Palmer L.E., de la Bastide M., Spiegel L., Nascimento L., Zutavern T., O'Shaughnessy A., Dike S., Dedhia N., Preston R., Balija V., McCombie W.R., Chow T.Y., Chen H.H., Chung M.C., Chen C.S., Shaw J.F., Wu H.P., Hsiao K.J., Chao Y.T., Chu M.K., Cheng C.H., Hour A.L., Lee P.F., Lin S.J., Lin Y.C., Liou J.Y., Liu S.M., Hsing Y.I., Raghuvanshi S., Mohanty A., Bharti A.K., Gaur A., Gupta V., Kumar D., Ravi V., Vij S., Kapur A., Khurana P., Khurana P., Khurana J.P., Tyagi A.K., Gaikwad K., Singh A., Dalal V., Srivastava S., Dixit A., Pal A.K., Ghazi I.A., Yadav M., Pandit A., Bhargava A., Sureshbabu K., Batra K., Sharma T.R., Mohapatra T., Singh N.K., Messing J., Nelson A.B., Fuks G., Kavchok S., Keizer G., Llaca E.L.V., Song R.T., Tanyolac B., Young S., Il K.H., Hahn J.H., Sangsakoo G., Vanavichit A., de Mattos L.A.T., Zimmer P.D., Malone G., Dellagostin O., de Oliveira A.C., Bevan M., Bancroft I., Minx P., Cordum H., Wilson R., Cheng Z.K., Jin W.W., Jiang J.M., Leong S.A., Iwama H., Gojobori T., Itoh T., Niimura Y., Fujii Y., Habara T., Sakai H., Sato Y., Wilson G., Kumar K., McCouch S., Juretic N., Hoen D., Wright S., 2005. The map-based sequence of the rice genome. Nature 436, 793-800.

Maury S., Geoffroy P., Legrand M., 1999. Tobacco *O*-methyltransferases involved in phenylpropanoid metabolism. The different caffeoyl-coenzyme A/5-hydroxyferuloyl-coenzyme A 3/5-O-methyltransferase and caffeic acid/5-hydroxyferulic acid 3/5-*O*-methyltransferase classes have distinct substrate specificities and expression patterns. Plant Physiol. 121, 215-223.

Mayer K., Schuller C., Wambutt R., Murphy G., Volckaert G., Pohl T., Dusterhoft A., Stiekema W., Entian K.D., Terryn N., Harris B., Ansorge W., Brandt P., Grivell L., Rieger M., Weichselgartner M., de Simone V., Obermaier B., Mache R., Muller M., Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidtheini T., Reichert B., Portatelle D., Perez-Alonso M., Boutry M., Bancroft I., Vos P., Hoheisel J., Zimmermann W., Wedler H., Ridley P., Langham S.A., McCullagh B., Bilham L., Robben J., Van der Schueren J., Grymonprez B., Chuang Y.J., Vandenbussche F., Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E., Weitzenegger T., Bothe G., Ramsperger U., Hilbert H., Braun M., Holzer E., Brandt A., Peters S., van

Staveren M., Dirkse W., Mooijman P., Lankhorst R.K., Rose M., Hauf J., Kotter P., Berneiser S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H., De Keyser A., Buysshaert C., Gielen J., Villarroel R., De Clercq R., Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S., Clark L., Doggett J., Hall S., Kay M., Lennard N., Mclay K., Mayes R., Pettett A., Rajandream M.A., Lyne M., Benes V., Rechmann S., Borkova D., Blocker H., Scharfe M., Grimm M., Lohnert T.H., Dose S., de Haan M., Maarse A., Schafer M., Muller-Auer S., Gabel C., Fuchs M., Fartmann B., Granderath K., Dauner D., Herzl A., Neumann S., Argiriou A., Vitale D., Liguori R., Piravandi E., Massenet O., Quigley F., Clabauld G., Mundlein A., Felber R., Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S., Chefdor F., Cooke R., Berger C., Montfort A., Casacuberta E., Gibbons T., Weber N., Vandenbol M., Bargues M., Terol J., Torres A., Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T., Heijnen L., Schwarz S., Scholler P., Heber S., Francs P., Bielke C., Frishman D., Haase D., Lemcke K., Mewes H.W., Stocker S., Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K., Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L., Sehkon M., Murray J., Sheet P., Cordes M., bu-Threideh J., Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J., Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D., Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K., Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L., Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D., Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshu C., Antonoiu B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C., Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R., Swaby I., 1999. Sequence and analysis of chromosome 4 of the plant *Arabidopsis thaliana*. Nature 402, 761-768.

Mederacke H., Biehl B., Selmar D., 1996. Characterization of two cyano glucosyltransferases from cassava leaves. Phytochemistry 42, 1517-1522.

Mederacke H., Selmar D., Biehl B., 1995. Glucosyltransferases of the cyanogenic plant cassava (*Manihot esculenta*, Crantz). J. Appl. Bot. Food Qual. 69, 119-124.

Meotti F.C., Ardenghi J.V., Pretto J.B., Souza W.M., Moura J.D.A., Cunha A., Soldi C., Pizzolatti M.G., Santos A.R.S., 2006. Antinociceptive properties of coumarins, steroid and dihydrostyryl-2-pyrones from *Polygala sabulosa* (Polygalaceae) in mice. J. Pharm. Pharmacol. 58, 107-112.

Meyermans H., Morreel K., Lapierre C., Pollet B., De Bruyn A., Busson R., Herdewijn P., Devreese B., Van Beeumen J., Marita J.M., Ralph J., Chen C.Y., Burggraeve B., Van Montagu M., Messens E., Boerjan W., 2000. Modifications in lignin and accumulation of phenolic glucosides in poplar xylem upon down-regulation of caffeoyl-coenzyme A *O*-methyltransferase, an enzyme involved in lignin biosynthesis. J. Biol. Chem. 275, 36899-36909.

Millar D.J., Long M., Donovan G., Fraser P.D., Boudet A.M., Danoun S., Bramley P.M., Bolwell G.P., 2007. Introduction of sense constructs of cinnamate 4-hydroxylase (CYP73A24) in transgenic tomato plants shows opposite effects on flux into stem lignin and fruit flavonoids. Phytochemistry 68, 1497-1509.

Mizutani M., Ward E., Dimaio J., Ohta D., Ryals J., Sato R., 1993. Molecular-cloning and sequencing of A cDNA-encoding mung bean cytochrome-P450 (P450C4H) possessing cinnamate 4-hydroxylase activity. Biochem. Biophys. Res. Commun. 190, 875-880.

Mohammadi M., Karr A.L., 2003. Induced lipoxygenases in soybean root nodules. Plant Sci. 164, 471-479.

Moon P.D., Lee B.H., Jeong H.J., An H.J., Park S.J., Kim H.R., Ko S.G., Um J.Y., Hong S.H., Kim H.M., 2007. Use of scopoletin to inhibit the production of inflammatory cytokines through inhibition of the I kappa B/NF-kappa B signal cascade in the human mast cell line HMC-1. Eur. J. Pharmacol. 555, 218-225.

Moriguchi K., Suzuki T., Ito Y., Yamazaki Y., Niwa Y., Kurata N., 2005. Functional isolation of novel nuclear proteins showing a variety of subnuclear localizations. Plant Cell 17, 389-403.

Nagashima S., Inagaki R., Kubo A., Hirotani M., Yoshikawa T., 2004. cDNA cloning and expression of isoflavonoid-specific glucosyltransferase from *Glycyrrhiza echinata* cell-suspension cultures. Planta 218, 456-459.

Nair P.M., Vining L.C., 1965. Cinnamic acid hydroxylase in spinach. Phytochemistry 4, 161-168.

Nakamura Y., Kaneko T., Sato S., Mimuro M., Miyashita H., Tsuchiya T., Sasamoto S., Watanabe A., Kawashima K., Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N., Shimpo S., Takeuchi C., Yamada M., Tabata S., 2003. Complete genome structure of *Gloeobacter violaceus* PCC 7421, a cyanobacterium that lacks thylakoids. DNA Res. 10, 137-145.

Nassar N., 2001. Cassava, *Manihot esculenta* Crantz and wild relatives: Their relationships and evolution. Genet. Resour. Crop Evol. 48, 429-436.

Nassar N.M.A., Hashimoto D.Y.C., Fernandes S.D.C., 2008. Wild Manihot species: botanical aspects, geographic distribution and economic value. Genet. Mol. Res. 7, 16-28.

Nedelkina S., Jupe S.C., Blee K.A., Schalk M., Werck-Reichhart D., Bolwell G.P., 1999. Novel characteristics and regulation of a divergent cinnamate 4-hydroxylase (CYP73A15) from French bean: engineering expression in yeast. Plant Mol. Biol. 39, 1079-1090.

Nikawa J., Hosaka K., 1995. Isolation and characterization of genes that promote the expression of inositol transporter gene itr1 in *Saccharomyces cerevisiae*. Mol. Microbiol. 16, 301-308.

Noon R.A., Booth R.H., 1977. Nature of post-harvest deterioration of cassava roots. Trans. Br. Mycol. Soc. 69, 287-290.

Ohyanagi H., Tanaka T., Sakai H., Shigemoto Y., Yamaguchi K., Habara T., Fujii Y., Antonio B.A., Nagamura Y., Imanishi T., Ikeo K., Itoh T., Gojobori T., Sasaki T., 2006. The rice annotation project database (RAP-DB): hub for *Oryza sativa* ssp *japonica* genome information. Nucleic Acids Res. 34, D741-D744.

Ojewole J.A.O., Adesina S.K., 1983. Mechanism of the hypotensive effect of scopoletin isolated from the fruit of *Tetrapleura tetraptera*. Planta Med. 49, 46-50.

Oliveira E.J., Romero M.A., Silva M.S., Silva B.A., Medeiros I.A., 2001. Intracellular calcium mobilization as a target for the spasmolytic action of scopoletin. Planta Med. 67, 605-608.

Olsen K.M., Schaal B.A., 1999. Evidence on the origin of cassava: Phylogeography of *Manihot esculenta*. Proc. Natl. Acad. Sci. U. S. A. 96, 5586-5591.

Olson M.M., Roseland C.R., 1991. Induction of the coumarins scopoletin and ayapin in sunflower by insect feeding stress and effects of coumarins on the feeding of sunflower beetle (Coleoptera, Chrysomelidae). Environ. Entomol. 20, 1166-1172.

Onyesom I., Okoh P.N., 2006. Quantitative analysis of nitrate and nitrite contents in vegetables commonly consumed in Delta State, Nigeria. Br. J. Nutr. 96, 902-905.

Orgambide G.G., Hollingsworth R.I., Dazzo F.B., 1992. Structural characterization of a novel diglycosyl diacylglyceride glycolipid from *Rhizobium trifolii* ANU843. Carbohydr. Res. 233, 151-159.

Page R.D.M., 1996. TreeView: An application to display phylogenetic trees on personal computers. Comput. Appl. Biosci. 12, 357-358.

Pallas J.A., Paiva N.L., Lamb C., Dixon R.A., 1996. Tobacco plants epigenetically suppressed in phenylalanine ammonia-lyase expression do not develop systemic acquired resistance in response to infection by tobacco mosaic virus. Plant J. 10, 281-293.

Panda S., Kar A., 2006. Evaluation of the antithyroid, antioxidative and antihyperglycemic activity of scopoletin from *Aegle marmelos* leaves in hyperthyroid rats. Phytother. Res. 20, 1103-1105.

Paquette S., Moller B.L., Bak S., 2003. On the origin of family 1 plant glycosyltransferases. Phytochemistry 62, 399-413.

Park C., Jin C.Y., Kim G.Y., Choi I.W., Kwon T.K., Choi B.T., Lee S.J., Lee W.H., Choi Y.H., 2008. Induction of apoptosis by esculetin in human leukemia U937 cells through activation of JNK and ERK. Toxicol. Appl. Pharmacol. 227, 219-228.

Peng L.C., Kawagoe Y., Hogan P., Delmer D., 2002. Sitosterol-β-glucoside as primer for cellulose synthesis in plants. Science 295, 147-150.

Pereira L.F., Goodwin P.H., Erickson L., 1999. The role of a phenylalanine ammonia lyase gene during cassava bacterial blight and cassava bacterial necrosis. J. Plant Res. 112, 51-60.

Petersen M, Strack D, Matern U, 1999. Biosynthesis of phenylpropanoids and related compounds. In M Wink, ed, Biochemistry of Plant Secondary Metabolism. Academic Press, Sheffield, pp 151-185.

Peterson J.K., Harrison H.F., Jackson D.M., Snook M.E., 2003. Biological activities and contents of scopolin and scopoletin in sweetpotato clones. HortScience 38, 1129-1133.

Pham T.N., Kover K.E., Jin L., Uhrin D., 2005. ^{1}H-detected double-*J*-modulated INEPT-INADEQUATE for simultaneous determination of one-bond and long-range carbon-carbon connectivities and the measurement of all carbon-carbon coupling constants. J. Magn. Reson. 176, 199-206.

Prats E., Bazzalo M.E., Leon A., Jorrin J.V., 2006. Fungitoxic effect of scopolin and related coumarins on *Sclerotinia sclerotiorum*. A way to overcome sunflower head rot. Euphytica 147, 451-460.

Prawat H., Mahidol C., Ruchirawat S., Prawat U., Tuntiwachwuttikul P., Tooptakong U., Taylor W.C., Pakawatchai C., Skelton B.W., White A.H., 1995. Cyanogenic and non-cyanogenic glycosides from *Manihot esculenta*. Phytochemistry 40, 1167-1173.

Rataboul P., Alibert G., Boller T., Boudet A.M., 1985. Intracellular transport and vacuolar accumulation of *o*-coumaric acid glucoside in *Melitolus alba* mesophyll cell protoplasts. Biochim. Biophys. Acta - Biomembranes 816, 25-36.

Rawel H.M., Kroll J., 2003. The importance of cassava (*Manihot esculenta* Crantz) as the main staple food in tropical countries. Dtsch. Lebensm. -Rundsch. 99, 102-108.

Reif B., Kock M., Kerssebaum R., Schleucher J., Griesinger C., 1996. Determination of ^{1}J, ^{2}J, and ^{3}J carbon-carbon coupling constants at natural abundance. J. Magn. Reson. B 112, 295-301.

Reilly K., Bernal D., Cortes D.F., Gómez-Vásquez R., Tohme J., Beeching J.R., 2007. Towards identifying the full set of genes expressed during cassava post-harvest physiological deterioration. Plant Mol. Biol. 64, 187-203.

Reilly K., Gómez-Vásquez R., Buschmann H., Tohme J., Beeching J.R., 2003. Oxidative stress responses during cassava post-harvest physiological deterioration. Plant Mol. Biol. 53, 669-685.

Reilly K., Gómez-Vásquez R., Buschmann H., Tohme J., Beeching J.R., 2004. Oxidative stress responses during cassava post-harvest physiological deterioration. Plant Mol. Biol. 56, 625-641.

Richert J., Kranz E., Lorz H., Dresselhaus T., 1996. A reverse transcriptase-polymerase chain reaction assay for gene expression studies at the single cell level. Plant Sci. 114, 93-99.

Riveiro M.E., Moglioni A., Vazquez R., Gomez N., Facorro G., Piehl L., de Celis E.R., Shayo C., Davio C., 2008. Structural insights into hydroxycoumarin-induced apoptosis in U-937 cells. Bioorg. Med. Chem. 16, 2665-2675.

Rizk A.M., Rimpler H., 1972. Isolation of daphnoretin and β-Sitosterol-β-D-glucoside from *Thymelea hirsuta*. Phytochemistry 11, 473-&.

Ro D.K., Mah N., Ellis B.E., Douglas C.J., 2001. Functional characterization and subcellular localization of poplar (*Populus trichocarpa* x *Populus deltoides*) cinnamate 4-hydroxylase. Plant Physiol. 126, 317-329.

Robbins R.J., Schmidt W.F., 2004. Optimized synthesis of four isotopically labeled (^{13}C-enriched) phenolic acids via a malonic acid condensation. J. Labelled Compd. Radiopharm. 47, 797-806.

Rogers D, Appan S, 1973. Flora Neotropica Monograph. Hafner Press, New York.

Rollinger J.M., Hornick A., Langer T., Stuppner H., Prast H., 2004. Acetylcholinesterase inhibitory activity of scopolin and scopoletin discovered by virtual screening of natural products. J. Med. Chem. 47, 6248-6254.

Roseland C.R., Espinasse A., Grosz T.J., 1991. Somaclonal variants of sunflower with modified coumarin expression under stress. Euphytica 54, 183-190.

Runeckles V.C., 1963. Tobacco polyphenols. III. The role of ferulic acid in the biosynthesis of chlorogenic acid. Can. J. Biol. Physiol. 41, 2259-2267.

Russell D.W., Conn E.E., 1967. The cinnamic acid 4-hydroxylase of pea seedlings. Arch. Biochem. Biophys. 122, 256-258.

Ryan W.L., Gordon D.J., Levy D.H., 2002. Gas-phase photochemistry of the photoactive yellow protein chromophore *trans-p*-coumaric acid. J. Am. Chem. Soc. 124, 6194-6201.

Sakai T., Nakagawa Y., 1988. Diterpenic stress metabolites from cassava roots. Phytochemistry 27, 3769-3779.

Sakai T., Nakagawa Y., Uritani I., Data E.S., 1986. Occurrence of various kinds of metabolites in physiologically and microbially damaged cassava (*Manihot esculenta* crantz) roots. Agric. Biol. Chem. 50, 2905-2907.

Sambrook J, Russell DW, 2001. Molecular cloning: a laboratory manual. Cold Spring Harbor, N.Y. : Cold Spring Harbor Laboratory.

Sanchez T., Chavez A.L., Ceballos H., Rodriguez-Amaya D., Nestel P., Ishitani M., 2006. Reduction or delay of post-harvest physiological deterioration in cassava roots with higher carotenoid content. J. Sci. Food. Agric. 86, 634-639.

Sato T., Kiuchi F., Sankawa U., 1982. Inhibition of phenylalanine ammonia-lyase by cinnamic acid derivatives and related compounds. Phytochemistry 21, 845-850.

Schmidt H.L., Werner R.A., Eisenreich W., Fuganti C., Fronza G., Remaud G., Robins R.J., 2006. The prediction of isotopic patterns in phenylpropanoids from their precursors and the mechanism of the NIH-shift: Basis of the isotopic characteristics of natural aromatic compounds. Phytochemistry 67, 1094-1103.

Schmitt D., Pakusch A.E., Matern U., 1991. Molecular-Cloning, Induction, and Taxonomic Distribution of Caffeoyl-Coa 3-*O*-Methyltransferase, An Enzyme Involved in Disease Resistance. J. Biol. Chem. 266, 17416-17423.

Schoch G., Goepfert S., Morant M., Hehn A., Meyer D., Ullmann P., Werck-Reichhart D., 2001. CYP98A3 from *Arabidopsis thaliana* is a 3 '-hydroxylase of phenolic esters, a missing link in the phenylpropanoid pathway. J. Biol. Chem. 276, 36566-36574.

Schröder N.W.J., Schombel U., Heine H., Göbel U.B., Zähringer U., Schumann R.R., 2003. Acylated cholesteryl galactoside as a novel immunogenic motif in *Borrelia burgdorferi sensu stricto*. J. Biol. Chem. 278, 33645-33653.

Seltzer S., Lin M., 1979. Maleylacetone *cis-trans*-isomerase - Mechanism of the interaction of coenzyme glutathione and substrate maleylacetone in the presence and absence of enzyme. J. Am. Chem. Soc. 101, 3091-3097.

Shah J., 2003. The salicylic acid loop in plant defense. Curr. Opin. Plant Biol. 6, 365-371.

Shin S.Y., Lee H.S., Kwon S.Y., Kwon S.T., Kwak S.S., 2005. Molecular characterization of a cDNA encoding copper/zinc superoxide dismutase from cultured cells of *Manihot esculenta*. Plant Physiol. Biochem. 43, 55-60.

Silvan A.M., Abad M.J., Bermejo P., Sollhuber M., Villar A., 1996. Antiinflammatory activity of coumarins from *Santolina oblongifolia*. J. Nat. Prod. 59, 1183-1185.

Siritunga D., Sayre R., 2004. Engineering cyanogen synthesis and turnover in cassava (*Manihot esculenta*). Plant Mol. Biol. 56, 661-669.

Siritunga D., Sayre R., 2007. Transgenic approaches for cyanogen reduction in cassava. Journal of Aoac International 90, 1450-1455.

Sirois J.C., Miller R.W., 1972. Mechanism of scopoletin-induced inhibition of peroxidase catalyzed degradation of indole-3-acetate. Plant Physiol. 49, 1012-&.

Snook M.E., Chortyk O.T., Csinos A.S., 1991. Black shank disease fungus - inhibition of growth by tobacco root constituents and related-compounds. ACS Symp. Ser. 449, 388-398.

Son D., Lee P., Lee J., Lee S., Choi S.Y., Lee J.W., Kim S.Y., 2007. Neuroprotective effect of scopoletin from *Angelica dahurica* on oxygen and glucose deprivation-exposed rat organotypic hippocampal slice culture. Food Sci. Biotechnol. 16, 632-635.

Stoker J.R., 1964. The biosynthesis of coumarin in *Melilotus alba*. Biochem. Biophys. Res. Commun. 14, 17-20.

Strack D, 1997. Phenolic metabolism. In PM Dey, JB Harborne, eds, Plant Biochemistry. Academic Press, London, pp 387-416.

Stupak M., Vanderschuren H., Gruissem W., Zhang P., 2006. Biotechnological approaches to cassava protein improvement. Trends Food Sci. Technol. 17, 634-641.

Swigonova Z., Lai J.S., Ma J.X., Ramakrishna W., Llaca V., Bennetzen J.L., Messing J., 2004. Close split of sorghum and maize genome progenitors. Genome Res. 14, 1916-1923.

Taguchi G., Imura H., Maeda Y., Kodaira R., Hayashida N., Shimosaka M., Okazaki M., 2000. Purification and characterization of UDP-glucose: hydroxycoumarin 7-O-glucosyltransferase, with broad substrate specificity from tobacco cultured cells. Plant Sci. 157, 105-112.

Taguchi G., Nakamura M., Hayashida N., Okazaki M., 2003a. Exogenously added naphthols induce three glucosyltransferases, and are accumulated as glucosides in tobacco cells. Plant Sci. 164, 231-240.

Taguchi G., Ubukata T., Hayashida N., Yamamoto H., Okazaki M., 2003b. Cloning and characterization of a glucosyltransferase that reacts on 7-hydroxyl group of flavonol and 3-hydroxyl group of coumarin from tobacco cells. Arch. Biochem. Biophys. 420, 95-102.

Taguchi G., Yazawa T., Hayashida N., Okazaki M., 2001. Molecular cloning and heterologous expression of novel glucosyltransferases from tobacco cultured cells that have broad substrate specificity and are induced by salicylic acid and auxin. Eur. J. Biochem. 268, 4086-4094.

Tanaka Y., Data E.S., Hirose S., Taniguchi T., Uritani I., 1983. Biochemical-changes in secondary metabolites in wounded and deteriorated cassava roots. Agric. Biol. Chem. 47, 693-700.

Tanaka Y., Yonekura K., FukuchiMizutani M., Fukui Y., Fujiwara H., Ashikari T., Kusumi T., 1996. Molecular and biochemical characterization of three anthocyanin synthetic enzymes from *Gentiana triflora*. Plant Cell Physiol. 37, 711-716.

Taylor N., Chavarriaga P., Raemakers K., Siritunga D., Zhang P., 2004. Development and application of transgenic technologies in cassava. Plant Mol. Biol. 56, 671-688.

Teutsch H.G., Hasenfratz M.P., Lesot A., Stoltz C., Garnier J.M., Jeltsch J.M., Durst F., Werckreichhart D., 1993. Isolation and sequence of cDNA-encoding the jerusalem-artichoke cinnamate 4-hydroxylase, a major plant cytochrome-P450 involved in the general phenylpropanoid pathway. Proc. Natl. Acad. Sci. U. S. A. 90, 4102-4106.

Thorsoe K.S., Bak S., Olsen C.E., Imberty A., Breton C., Moller B.L., 2005. Determination of catalytic key amino acids and UDP sugar donor specificity of the cyanohydrin glycosyltransferase UGT85B1 from *Sorghum bicolor*. Molecular modeling substantiated by site-specific mutagenesis and biochemical analyses. Plant Physiol. 139, 664-673.

Tomkins J., Fregene M., Main D., Kim H., Wing R., Tohme J., 2004. Bacterial artificial chromosome (BAC) library resource for positional cloning of pest and disease resistance genes in cassava (*Manihot esculenta* Crantz). Plant Mol. Biol. 56, 555-561.

Tulloch A.P., Mazurek M., 1973. Deuterium isotope effects on ^{13}C chemical shifts in long chain aliphatic compounds. J. Chem. Soc. Chem. Commun. 692-693.

Turner L.B., Muellerharvey I., Mcallan A.B., 1993. Light-induced isomerization and dimerization of cinnamic acid-derivatives in cell-walls. Phytochemistry 33, 791-796.

Valle T., Lopez J.L., Hernandez J.M., Corchete P., 1997. Antifungal activity of scopoletin and its differential accumulation in *Ulmus pumila* and *Ulmus campestris* cell suspension cultures infected with *Ophiostoma ulmi* spores. Plant Sci. 125, 97-101.

Verdier V., Restrepo S., Mosquera G., Jorge V., Lopez C., 2004. Recent progress in the characterization of molecular determinants in the *Xanthomonas axonopodis* pv. manihotis- cassava interaction. Plant Mol. Biol. 56, 573-584.

Vogt T., Grimm R., Strack D., 1999. Cloning and expression of a cDNA encoding betanidin 5-O-glucosyltransferase, a betanidin- and flavonoid-specific enzyme with high homology to inducible glucosyltransferases from the Solanaceae. Plant J. 19, 509-519.

Vogt T., Jones P., 2000. Glycosyltransferases in plant natural product synthesis: characterization of a supergene family. Trends Plant Sci. 5, 380-386.

Walton N.J., Mayer M.J., Narbad A., 2003. Molecules of interest - Vanillin. Phytochemistry 63, 505-515.

Wang S.H., Kuo S.C., Chen S.C., 2003. High-performance liquid chromatography determination of methionine adenosyltransferase activity using catechol-*O*-methyltransferase-coupled fluorometric detection. Anal. Biochem. 319, 13-20.

Watanabe K., Ito A., Sato T., Saito T., Hayashi H., Niitani Y., 1999. Esculetin suppresses proteoglycan metabolism by inhibiting the production of matrix metalloproteinases in rabbit chondrocytes. Eur. J. Pharmacol. 370, 297-305.

Watts K.T., Lee P.C., Schmidt-Dannert C., 2004. Exploring recombinant flavonoid biosynthesis in metabolically engineered *Escherichia coli*. ChemBioChem 5, 500-507.

Wehrli FW, Wirthlin T, 1983. Interpretation of ^{13}C NMR Spectra. John Wiley and Sons, London, p 108.

Welch R.M., 2002. Breeding strategies for biofortified staple plant foods to reduce micronutrient malnutrition globally. J. Nutr. 132, 495S-499S.

Wheatley C.C., Schwabe W.W., 1985. Scopoletin involvement in post-harvest physiological deterioration of cassava root (*Manihot esculenta* Crantz). J. Exp. Bot. 36, 783-791.

Williams D, Fleming I, 2008b. Spectroscopic Methods in Organic Chemistry. McGraw-Hill, Maidenhead, p 110.

Williams D, Fleming I, 2008a. Spectroscopic Methods in Organic Chemistry. McGraw-Hill, Maidenhead, pp 162-163.

Willits M.G., Giovanni M., Prata R.T.N., Kramer C.M., De Luca V., Steffens J.C., Graser G., 2004. Bio-fermentation of modified flavonoids: an example of in vivo diversification of secondary metabolites. Phytochemistry 65, 31-41.

Wise R.P., Rohde W., Salamini F., 1990. Nucleotide-sequence of the bronze-1 homologous gene from *Hordeum vulgare*. Plant Mol. Biol. 14, 277-279.

Wong W.S., Guo D., Wang X.L., Yin Z.Q., Xia B., Li N., 2005. Study of *cis*-cinnamic acid in *Arabidopsis thaliana*. Plant Physiol. Biochem. 43, 929-937.

Wu C.R., Huang M.Y., Lin Y.T., Ju H.Y., Ching H., 2007. Antioxidant properties of Cortex Fraxini and its simple coumarins. Food Chem. 104, 1464-1471.

Yamazaki M., Gong Z., Fukuchi-Mizutani M., Fukui Y., Tanaka Y., Kusumi T., Saito K., 1999. Molecular cloning and biochemical characterization of a novel anthocyanin 5-*O*-glucosyltransferase by mRNA differential display for plant forms regarding anthocyanin. J. Biol. Chem. 274, 7405-7411.

Yang D.H., Chung B.Y., Kim J.S., Kim J.H., Yun P.Y., Lee Y.K., Lim Y.P., Lee M.C., 2005. cDNA cloning and sequence analysis of the rice cinnamate-4'-hydroxylase gene, a cytochrome P450-dependent monooxygenase involved in the general phenylpropanoid pathway. J. Plant Biol. 48, 311-318.

Yang J.Y., Koo J.H., Yoon H.Y., Lee J.H., Park B.H., Kim J.S., Chi M.S., Park J.W., 2007. Effect of scopoletin on lipoprotein lipase activity in 3T3-L1 adipocytes. International J Mol. Med. 20, 527-531.

Zhang K., Yang E.B., Tang W.Y., Wong K.P., Mack P., 1997. Inhibition of glutathione reductase by plant polyphenols. Biochem. Pharmacol. 54, 1047-1053.

Zhang P., Vanderschuren H., Futterer J., Gruissem W., 2005. Resistance to cassava mosaic disease in transgenic cassava expressing antisense RNAs targeting virus replication genes. Plant Biotechnol. J. 3, 385-397.

Zhang S.H., Yang Q., Ma R.C., 2007. *Erwinia carotovora* ssp. *carotovora* infection induced "defense lignin" accumulation and lignin biosynthetic gene expression in chinese cabbage (*Brassica rapa* L. ssp *pekinensis*). J. Integr. Plant Biol. 49, 993-1002.

Zheng A.L., Wang W., Zhang H.J., Wang B., 1999. Two new improved approaches to the synthesis of coumarin-based prodrugs. Tetrahedron 55, 4237-4254.

Appendix 1
X-ray crystal data

X-Ray Crystallography

A single crystal of the compound ferulic -2-^{13}C acid was analyzed on a Nonius-kappaCCD diffractometer, using graphite monochromated Mo-K_a radiation of wavelength 0.71073Å. The sample was maintained at 150 K throughout the data collection, using an Oxford Cryosystems cryostat.

Details of the data collection and refinement are given in Tables 1-5.

Table 1
Crystal data and structure refinement for 1.

Identification code	k07ib1
Empirical formula	C10 H10 O4
Formula weight	194.18
Temperature	150(2) K
Wavelength	0.71073 Å
Crystal system	Monoclinic
Space group	P21/n
Unit cell dimensions	a = 4.6060(1)Å alpha = 90°
	b = 16.7630(4)Å beta = 91.408(1)°
	c = 11.8460(3)Å gamma = 90°
Volume	914.36(4) Å3
Z	4
Density (calculated)	1.411 Mg/m^3
Absorption coefficient	0.110 mm^{-1}
F(000)	408
Crystal size	0.30 x 0.25 x 0.25 mm
Theta range for data collection	3.65 to 27.50°
Index ranges	-5<=h<=5; -21<=k<=21; -15<=l<=15
Reflections collected	16061
Independent reflections	2083 [R(int) = 0.0389]
Reflections observed (>2sigma)	1819
Data Completeness	0.994
Absorption correction	None
Refinement method	Full-matrix least-squares on F^2
Data / restraints / parameters	2083 / 0 / 132
Goodness-of-fit on F2	1.061
Final R indices [I>2sigma(I)]	R1 = 0.0387 wR2 = 0.0995
R indices (all data)	R1 = 0.0463 wR2 = 0.1056
Largest diff. peak and hole	0.219 and -0.273 eÅ^{-3}

Hydrogen bonds with H..A < r(A) + 2.000 Angstroms and <DHA > 110 deg.

D-H	d(D-H)	d(H..A)	<DHA	d(D..A)	A
O2-H2A	0.840	1.790	176.34	2.629	O1 [-x-2, -y+1, -z+2]
O4-H4	0.840	2.088	152.38	2.859	O2 [x+3/2, -y+1/2, z-1/2]
O4-H4	0.840	2.199	114.49	2.659	O3

Table 2
Atomic coordinates (x 10^4) and equivalent isotropic displacement parameters ($Å^2$ x 10^3) for 1.U(eq) is defined as one third of the trace of the orthogonalized Uij tensor.

Atom	x	Y	z	U(eq)
O(1)	-7638(2)	5032(1)	8995(1)	30(1)
C(1)	-7128(2)	4334(1)	9318(1)	23(1)
O(2)	-8453(2)	4009(1)	10161(1)	29(1)
C(2)	-4988(3)	3834(1)	8757(1)	25(1)
O(3)	2631(2)	1808(1)	6902(1)	31(1)
C(3)	-3611(2)	4109(1)	7856(1)	24(1)
O(4)	4615(2)	2612(1)	5145(1)	29(1)
C(4)	1597(4)	1360(1)	7833(1)	38(1)
C(1')	-1486(2)	3697(1)	7171(1)	24(1)
C(2')	-541(3)	2916(1)	7413(1)	24(1)
C(3')	1493(3)	2554(1)	6744(1)	23(1)
C(4')	2618(3)	2960(1)	5816(1)	24(1)
C(5')	1696(3)	3728(1)	5574(1)	28(1)
C(6')	-338(3)	4093(1)	6247(1)	28(1)

Table 3
Bond lengths [Å] and angles [°] for 1

O(1)-C(1)	1.2510(14)	C(1)-O(2)	1.3028(14)
C(1)-C(2)	1.4652(15)	O(2)-H(2A)	0.8400
C(2)-C(3)	1.3367(17)	C(2)-H(2)	0.9500
O(3)-C(3')	1.3673(14)	O(3)-C(4)	1.4262(15)
C(3)-C(1')	1.4606(15)	C(3)-H(3)	0.9500
O(4)-C(4')	1.3610(14)	O(4)-H(4)	0.8400
C(4)-H(4A)	0.9800	C(4)-H(4B)	0.9800
C(4)-H(4C)	0.9800	C(1')-C(6')	1.3948(17)
C(1')-C(2')	1.4067(16)	C(2')-C(3')	1.3816(16)
C(2')-H(2')	0.9500	C(3')-C(4')	1.4030(16)
C(4')-C(5')	1.3842(17)	C(5')-C(6')	1.3871(17)
C(5')-H(5')	0.9500	C(6')-H(6')	0.9500
O(1)-C(1)-O(2)	122.48(10)	O(1)-C(1)-C(2)	121.34(10)
O(2)-C(1)-C(2)	116.18(10)	C(1)-O(2)-H(2A)	109.5
C(3)-C(2)-C(1)	120.17(11)	C(3)-C(2)-H(2)	119.9
C(1)-C(2)-H(2)	119.9	C(3')-O(3)-C(4)	116.94(9)
C(2)-C(3)-C(1')	128.21(11)	C(2)-C(3)-H(3)	115.9
C(1')-C(3)-H(3)	115.9	C(4')-O(4)-H(4)	109.5
O(3)-C(4)-H(4A)	109.5	O(3)-C(4)-H(4B)	109.5
H(4A)-C(4)-H(4B)	109.5	O(3)-C(4)-H(4C)	109.5
H(4A)-C(4)-H(4C)	109.5	H(4B)-C(4)-H(4C)	109.5
C(6')-C(1')-C(2')	118.76(11)	C(6')-C(1')-C(3)	118.88(11)
C(2')-C(1')-C(3)	122.36(10)	C(3')-C(2')-C(1')	120.15(11)
C(3')-C(2')-H(2')	119.9	C(1')-C(2')-H(2')	119.9
O(3)-C(3')-C(2')	125.88(11)	O(3)-C(3')-C(4')	113.75(10)

C(2')-C(3')-C(4')	120.36(11)	O(4)-C(4')-C(5')	119.11(10)
O(4)-C(4')-C(3')	121.16(11)	C(5')-C(4')-C(3')	119.73(11)
C(4')-C(5')-C(6')	119.92(11)	C(4')-C(5')-H(5')	120.0
C(6')-C(5')-H(5')	120.0	C(5')-C(6')-C(1')	121.07(11)
C(5')-C(6')-H(6')	119.5	C(1')-C(6')-H(6')	119.5

Symmetry transformations used to generate equivalent atoms:

Table 4
Anisotropic displacement parameters ($Å^2$ x 10^3) for 1. The anisotropic displacement factor exponent takes the form: -2 gpi2 [h2 a*2 U11 + ... + 2 h k a* b* U

Atom	U11	U22	U33	U23	U13	U12
O(1)	34(1)	23(1)	33(1)	-1(1)	9(1)	4(1)
C(1)	23(1)	22(1)	24(1)	-3(1)	3(1)	-2(1)
O(2)	28(1)	26(1)	32(1)	-2(1)	12(1)	0(1)
C(2)	23(1)	22(1)	28(1)	-3(1)	3(1)	1(1)
O(3)	39(1)	23(1)	30(1)	1(1)	13(1)	6(1)
C(3)	23(1)	23(1)	28(1)	-3(1)	3(1)	1(1)
O(4)	34(1)	28(1)	26(1)	-1(1)	12(1)	3(1)
C(4)	58(1)	25(1)	32(1)	4(1)	16(1)	7(1)
C(1')	22(1)	25(1)	24(1)	-3(1)	3(1)	0(1)
C(2')	24(1)	25(1)	22(1)	-1(1)	4(1)	-2(1)
C(3')	25(1)	22(1)	23(1)	-2(1)	2(1)	0(1)
C(4')	24(1)	27(1)	21(1)	-4(1)	5(1)	0(1)
C(5')	31(1)	29(1)	25(1)	3(1)	8(1)	1(1)
C(6')	31(1)	25(1)	29(1)	3(1)	4(1)	3(1)

Table 5.
Hydrogen coordinates (x 10^4) and isotropic displacement, parameters ($Å^2$ x 10^3) for 1.

Atom	x	Y	z	U(eq)
H(2A)	-9678	4331	10411	34
H(2)	-4578	3314	9035	30
H(3)	-4074	4639	7633	29
H(4)	4954	2146	5374	60(6)
H(4A)	2043	1647	8537	57
H(4B)	2546	837	7854	57
H(4C)	-509	1289	7748	57
H(2')	-1303	2636	8037	28
H(5')	2455	4006	4947	34
H(6')	-960	4620	6077	34

Appendix 2

Abstracts and publications arising from these studies

Abstract of an oral entitled: "Secondary metabolite production and genetic analysis in cassava"

Soad A. L. Bayoumi, Michael G. Rowan, Ian S. Blagbrough, and John R Beeching*
Department of Pharmacy and Pharmacology, University of Bath, Bath BA2 7AY, and
*Department of Biology and Biochemistry, University of Bath, Bath BA2 7AY, U.K.
prsisb@bath.ac.uk, bssjrb@bath.ac.uk

The Young Scientists Symposium "Future trends in phytochemistry" Olomouc, 27th June to 1st July 2006

Abstract of a poster entitled: "Phytochemical analysis of cassava roots during post-harvest physiological deterioration"

Soad A. L. Bayoumi, Michael G. Rowan, Ian S. Blagbrough, and John R Beeching*
Department of Pharmacy and Pharmacology, University of Bath, Bath BA2 7AY, and
*Department of Biology and Biochemistry, University of Bath, Bath BA2 7AY, U.K.
prsisb@bath.ac.uk, bssjrb@bath.ac.uk

The 50 years of the phytochemical society of Europe "Highlights in the evolution of phytochemistry" Cambridge, 11th to 14th April 2007

Abstract of an oral entitled: "A molecular and genetic search for coumarin glucosyltransferases in cassava roots"

Soad A. L. Bayoumi, Michael G. Rowan, Ian S. Blagbrough, and John R Beeching*
Department of Pharmacy and Pharmacology, University of Bath, Bath BA2 7AY, and
*Department of Biology and Biochemistry, University of Bath, Bath BA2 7AY, U.K.
prsisb@bath.ac.uk, bssjrb@bath.ac.uk

The Young Scientists Symposium "Future trends in phytochemistry" Gargnano, 6th to 9th June 2007

Abstract of an oral entitled: "Biosynthesis of scopoletin and scopolin in cassava roots during post-harvest physiological deterioration"

Soad A. L. Bayoumi, Michael G. Rowan, Ian S. Blagbrough, and John R Beeching*
Department of Pharmacy and Pharmacology, University of Bath, Bath BA2 7AY, and
*Department of Biology and Biochemistry, University of Bath, Bath BA2 7AY, U.K.
prsisb@bath.ac.uk, bssjrb@bath.ac.uk

The PSE Young Scientists' Meeting Future trends in phytochemistry Compounds – Enzymes -Genes Bad Herrenalb, 26th to 29th March 2008

Abstract of an oral entitled: "Isotopic labelling studies on the scopoletin biosynthetic pathways in cassava roots during post-harvest physiological deterioration"

Soad A. L. Bayoumi, Michael G. Rowan, Ian S. Blagbrough, and John R Beeching*
Department of Pharmacy and Pharmacology, University of Bath, Bath BA2 7AY, and
*Department of Biology and Biochemistry, University of Bath, Bath BA2 7AY, U.K.
prsisb@bath.ac.uk, bssjrb@bath.ac.uk

The RSC Postgraduate Symposium: Bioorganic Chemistry and Chemical Biology University of Bath, 31st March 2008

Abstract of an oral entitled: "Investigation of scopoletin biosynthesis during post-harvest physiological deterioration in cassava roots using stable isotopic labelling"

Soad A. L. Bayoumi, Michael G. Rowan, Ian S. Blagbrough, and John R Beeching*
Department of Pharmacy and Pharmacology, University of Bath, Bath BA2 7AY, and
*Department of Biology and Biochemistry, University of Bath, Bath BA2 7AY, U.K.
prsisb@bath.ac.uk, bssjrb@bath.ac.uk

The BPC Pharmacy in the 1st century – Adding years to life and life to years Manchester, 7th-9th September 2008

Secondary Metabolite Production and Genetic Analysis in Cassava

Soad A. L. Bayoumi, Michael G. Rowan, Ian S. Blagbrough, and John R Beeching*

Department of Pharmacy and Pharmacology, University of Bath, Bath BA2 7AY, and
*Department of Biology and Biochemistry, University of Bath, Bath BA2 7AY, U.K.

prsisb@bath.ac.uk, bssjrb@bath.ac.uk

Cassava (*Manihot esculenta*) tuberous roots are a staple food for more than five hundred million people in tropical countries due to their high starch content and ability to withstand unfavourable conditions. The crop suffers from some problems that limit its utilization such as it is susceptible to microbial and viral diseases, and its cyanogenic glycoside content. The release of HCN leads to neurological disorders, which are difficult-to-treat diseases like kunzo. In addition, the roots have a short shelf life due to post-harvest physiological deterioration (PPD). Within 2-3 days after harvesting the root shows blue to black vascular streaking and it becomes unpalatable and therefore unmarketable, this PPD significantly affects the crop's economic value. Biosynthetic studies show that within 15 min of the root being injured an oxidative burst occurs, which is followed by changes in the expression of genes and the accumulation of secondary metabolites (Sakai et al 1988; Reilly et al 2004). In order to more fully understand these phenomena, and with a view to overcoming this problem, we have analyzed cassava root extracts by chromatographic methods to isolate and then identify the active constituents by a variety of spectrometric techniques. We used both healthy roots and after PPD in order to look at secondary metabolite production.

Cassava roots were peeled; the outer 2-3 mm layer representing the phelloderm, was removed. The peeled roots were extracted with EtOH then CH_2Cl_2. β-carotene, glycoside of diacylglyceride and β-sitosterol glucoside were identified in the healthy root extract. While the other part of the roots were sliced then stored under controlled conditions (20 ^{o}C, 80-90% relative humidity, for 6 days), root slices were taken from the third day on a daily basis and peeled, crushed and from the third day then extracted with EtOH. Linamarin, scopolin, and scopoletin have been isolated from deteriorated root extract (HPLC) and identified by NMR spectroscopy.

Glucosyltransferases play important roles in many aspects of plant development and defence, in deteriorating cassava root the dominant coumarin, scopoletin is glycosylated to its glycoside scopolin. In order to identify the principle glucosyltransferases expressed in the deteriorating cassava root, glucosyltransferase nucleotide sequences from a range of plants were aligned and used to construct a phylogenetic tree of these enzymes. This tree was then used to identify clusters of related enzymes from which degenerate PCR primers were designed. These primers were used to amplify and clone glucosyltransferase DNAs from a cDNA library constructed from mRNA from deteriorated cassava roots. These cassava glucosyltransferases have been sequenced and characterised.

We acknowledge the financial support of the Egyptian Government (studentship to S.A.L.B).

Reilly, K., et al Oxidative stress responses during cassava post-harvest physiological deterioration. (2004) Plant Mol. Biol. **56**: 625-641
Sakai, T., Nakagawa, Y. Cassava diterpenic stress metabolites (1988) ***Phytochem*. 27:** 3769-3779

Phytochemical analysis of cassava roots during post-harvest physiological deterioration

Soad A. L. Bayoumi, Michael G. Rowan, Ian S. Blagbrough, and John R Beeching*

Department of Pharmacy and Pharmacology, University of Bath, Bath BA2 7AY, and *Department of Biology and Biochemistry, University of Bath, Bath BA2 7AY, U.K. prsisb@bath.ac.uk, bssjrb@bath.ac.uk

Cassava (*Manihot esculenta* Crantz Family Euphorbiaceae) is an important crop, the fourth most important food source in tropical countries due to its high root starch content. Although cassava is relatively easy to grow, even in poor soils and drought conditions, its roots are low in protein and contain cyanogenic glycosides which release HCN upon hydrolysis causing neurological disorders e.g. konzo in people who are malnourished. Also, cassava is susceptible to a variety of pests and diseases. These problems are being overcome by breeding programs and the production of transgenic cassava with low cyanide content and disease resistance. In addition, harvested roots have a very short shelf-life (of only one to three days) due to post-harvest physiological deterioration (PPD), which can cause significant wastage and economic losses. PPD was explained as a physiological process not due to microorganisms by Averre (1967) and by Noon and Booth (1977). Within 2-3 days of harvesting, the roots show blue to black vascular streaking and are unpalatable and therefore unmarketable, significantly affecting the crop's economic value. PPD has been explained on a molecular basis as an oxidative burst which occurs within 15 min of the root being injured (Reilly et al., 2003), followed by altered gene regulation and the accumulation of secondary metabolites, some of which show antioxidant properties such as the coumarin scopoletin and its glucoside scopolin. The aims of this project include a contribution to the understanding of secondary metabolite production and glucosyltransferase gene analysis during PPD in order to produce a new crop viable for the tropical regions of Africa and elsewhere. In order to complement and extend the range of glucosyltransferases known to be expressed during PPD, especially scopoletin glucosyltransferase, we are screening a PPD-related cDNA library by PCR and nucleic acid hybridization using amplified gene-probes with different established biological functions.

We acknowledge the financial support of the Egyptian Government (studentship to S.A.L.B).

References

Averre, C.W., 1967. Vascular streaking of stored cassava roots. Proceedings of the 1st International Symposium on Tropical Root Crops, Trinidad. 2, 31.

Noon, R.A., Booth, R.H., 1977. Nature of post-harvest deterioration of cassava roots. Trans. Br. Mycol. Soc. 69, 287-290.

Reilly, K., Gómez-Vásquez, R., Buschmann, H., Tohme, J., Beeching, J.R., 2003. Oxidative stress responses during cassava post-harvest physiological deterioration. Plant Molecular Biology. 53, 669-685.

A molecular and genetic search for coumarin glucosyltransferases in cassava roots

Soad A. L. Bayoumi, Michael G. Rowan, Ian S. Blagbrough, and John R Beeching*

Department of Pharmacy and Pharmacology, University of Bath, Bath BA2 7AY, and
*Department of Biology and Biochemistry, University of Bath, Bath BA2 7AY, U.K.
prsisb@bath.ac.uk, bssjrb@bath.ac.uk

Cassava (*Manihot esculenta* Crantz Family Euphorbiaceae) is the fourth most important food source in tropical countries due to its high root starch content. Cassava is easy to grow, even in poor soils and under drought conditions, but harvested roots have a short shelf-life (only up to 3 days) due to post-harvest physiological deterioration (PPD), a physiological process not due to microorganisms. The roots deteriorate with blue to black vascular streaking, becoming unpalatable and unmarketable; thus, PPD causes significant economic loss of cassava roots. PPD has been described as a reactive oxygen species mediated programmed cell death triggered by an oxidative burst which occurs within 15 min of the root being injured (Reilly et al., 2007), which includes altered gene regulation and the accumulation of secondary metabolites, e.g. the antioxidant coumarin scopoletin and its glucoside scopolin. The aims of this project include understanding secondary metabolite production and glucosyltransferase gene analysis during PPD in order to produce a viable new crop.

Peeled roots of cassava were sequentially extracted with EtOH then CH_2Cl_2. β-Carotene, β-sitosterol glucoside, and a diacylglyceride glycoside were identified from healthy roots. Another root sample was sliced then stored under controlled conditions (20 ^{o}C, 80-90% relative humidity, for 6 days). The sliced roots were taken on a daily basis (from the third day onwards), peeled and extracted with EtOH. Linamarin, scopolin, and its aglycone scopoletin were isolated from the deteriorated root extract by HPLC, and unambiguously identified by NMR spectroscopy. The isolation of glycosidic compounds prompted an investigation of glucosyltransferases as these enzymes play important roles in many aspects of plant development and defence. Glucosyltransferase amino acid sequences from a range of plant species: *M. esculenta*, *Nicotiana tabacum*, and *Arabidopsis thaliana*, were aligned (Clustal X) and used to generate a phylogenetic tree. Four distinct cluster groups of related glucosyltransferases were chosen in the tree, and nucleotide sequence alignments for each were carried out in order to design four pairs of forward and reverse degenerate PCR primers. These primers were used to amplify and clone glucosyltransferase DNAs from a cDNA library constructed from the mRNA of deteriorated cassava roots. Two glucosyltransferases have been identified. In order to extend the known range of glucosyltransferases expressed during PPD, and to identify the scopoletin glucosyltransferase, we are screening a PPD-related cDNA library by nucleic acid hybridization using probes with different established glucosyltransferase functions: tobacco scopoletin glucosyltransferase, strawberry cinnamate glucosyltransferase, and perilla anthocyanin glucosyltransferase. A new glucosyltransferase has been identified in cassava showing some similarity to scopoletin glucosyltransferase in tobacco.

We acknowledge financial support from the Egyptian Government (studentship to SALB).

Reilly, K., Bernal, D., Cortés, D.F., Gómez-Vásquez, R., Tohme, J., Beeching, J.R.
Plant Molecular Biology. 2007, DOI 10.1023/A:1020695719706 in press.

SL15 BIOSYNTHESIS OF SCOPOLETIN AND SCOPOLIN IN CASSAVA ROOTS DURING POST-HARVEST PHYSIOLOGICAL DETERIORATION

Soad A. L. Bayoumi, Michael G. Rowan, Ian S. Blagbrough, and John R. Beeching*
Department of Pharmacy and Pharmacology, University of Bath, Bath BA2 7AY, and
*Department of Biology and Biochemistry, University of Bath, Bath BA2 7AY, U.K.
salhb20@bath.ac.uk, prsmgr@bath.ac.uk, prsisb@bath.ac.uk, bssjrb@bath.ac.uk

Cassava (*Manihot esculenta* Crantz Family Euphorbiaceae) is an important crop due to its high root starch content. However, harvested cassava roots have a short shelf-life of only 1-3 days due to post-harvest physiological deterioration (PPD) causing significant wastage and economic loss. PPD is a physiological process not due to microorganisms. Within only 1-3 days of harvesting, the roots show blue to black vascular streaking and are unpalatable and unmarketable. PPD has been explained on a molecular basis as an oxidative burst occurring within 15 min of the root being injured, followed by altered gene expression and the accumulation of secondary metabolites, e.g. the antioxidant hydroxycoumarins scopoletin and esculetin, and their respective glucosides scopolin and esculin (Buschmann et al. 2000), important for plants and with many therapeutic actions. However, their biosynthetic pathways are not resolved (Petersen et al. 1999).

Cassava roots (cv Mcol22), within 2 h of harvesting, were fed with different deuterium labelled intermediates on the postulated biosynthetic pathway of scopoletin (e.g. *trans*-cinnamic-d_7, *trans*-cinnamic-β,2,3,4,5,6-d_6 and *trans*-cinnamic-2,3,4,5,6-d_5 acids) and PPD was allowed to occur. Ethanolic extracts of these deteriorated roots were separated (HPLC) and analysed by HRESI-MS. Deuteriated cinnamic acids were incorporated, and typically 4.6% of the scopoletin was deuteriated. Incorporation (in both scopoletin and scopolin) of only 3 deuterons strongly supports our thesis that the *trans*-*cis*-isomerization step is enzymatic, and not photochemical, as proposed in other plants, where 4 deuterons would be found (Edwards and Stoker 1968; Dugave and Demange 2003). There are three hypothetical pathways for the biosynthesis of scopoletin (Kai et al. 2006) via: 2,4-dihydroxycinnamate, caffeate, or ferulate, and so to investigate further the pathway of scopoletin biosynthesis during PPD in cassava roots competition feeding experiments were performed with these possible intermediates and deuteriated *trans*-cinnamic acids. If any of these unlabelled intermediates lie on the scopoletin biosynthetic pathway they should compete (if incorporation is possible) with the labelled cinnamate and therefore decrease the amount of labelled scopoletin produced. A typical reduction to 3.4% was found when unlabelled *trans*-2,4-dihydroxy-cinnamate was used, and no reduction with either caffeate or ferulate. We conclude that the major pathway for scopoletin and scopolin biosynthesis in cassava during PPD is through *trans*-2,4-dihydroxy-cinnamate, enzymatically isomerised and then ring closed to umbelliferone.

We acknowledge the financial support of the Egyptian Government (studentship to SALB).

Buschmann H, Rodriguez M X, Tohme J, Beeching J R (2000) Ann Bot 86: 1153-1160
Dugave C, Demange L (2003) Chem Rev 103: 2475-2532
Edwards K G, Stoker J R (1968) Phytochemistry 7: 73-77
Kai K, Shimizu B, Mizutani M, Watanabe K, Sakata K (2006) Phytochemistry 67: 379-386
Petersen M, Strack D, Matern U (1999) (Ed Wink M, Sheffield Acad Press) Ann Plant Rev 2: 161-167

Isotopic labelling studies on the scopoletin biosynthetic pathways in cassava roots during post-harvest physiological deterioration

Soad A. L. Bayoumi, Michael G. Rowan, Ian S. Blagbrough, and John R. Beeching*
Department of Pharmacy and Pharmacology, University of Bath, Bath BA2 7AY, U.K. and
*Department of Biology and Biochemistry, University of Bath, Bath BA2 7AY, U.K.
salhb20@bath.ac.uk, prsmgr@bath.ac.uk, prsisb@bath.ac.uk, bssjrb@bath.ac.uk

Cassava (*Manihot esculenta* Crantz) is an important crop due to its nutritional value. However, harvested cassava roots have a short shelf-life due to post-harvest physiological deterioration (PPD) causing significant wastage and economic loss. Hydroxycoumarins play important roles in plant development and defence; additionally, some are pharmacologically active. Despite their importance, key aspects of the biosynthesis of these secondary metabolites remain unresolved.[1] Here we exploit the increase in scopoletin accumulation in cassava roots [2] post-harvest to test alternative pathways for the biosynthesis of this hydroxycoumarin. So, within 2 h of harvesting, Cassava roots (cv Mcol22) were treated with aq. solutions (4% Na_2CO_3, pH = 7.5) of deuterium labelled intermediates on the postulated biosynthetic pathway of scopoletin (e.g. *trans*-cinnamic-d_7, *trans*-cinnamic-3,2',3',4',5',6'-d_6, *trans*-cinnamic-2',3',4',5',6'-d_5 and *trans*-cinnamic-2-d acids) and PPD was allowed to occur. Ethanolic extracts of the deteriorated roots were separated (HPLC) and analysed (HR ESI-MS). Deuteriated cinnamic acids were incorporated, and typically 29% of the scopoletin was deuteriated. Incorporation (in both scopoletin and scopolin) of only 3 deuterons when fed with *trans*-cinnamic-d_7 or -d_6 acid, and isolation of non-deuteriated scopoletin when cinnamic-2-d acid was fed, strongly supports our thesis that the *trans*-*cis*-isomerisation step is enzymatic, and not photochemical as found in other plants (where 4 deuterons are incorporated).

There are three hypothetical pathways for the biosynthesis of scopoletin [3] via: 2',4'-dihydroxy-cinnamate, caffeate, or ferulate. In order to investigate further the possible pathways of scopoletin biosynthesis during PPD in cassava roots, competition feeding experiments were performed with these possible intermediates and deuteriated *trans*-cinnamic acids. If any of these unlabelled intermediates lie on the scopoletin biosynthetic pathway they should compete (if incorporation is possible) with the labelled cinnamate and therefore decrease the amount of labelled scopoletin produced. A typical reduction (4.6 to 3.4%) was found when unlabelled *trans*-2',4'-dihydroxy-cinnamate was used, but no reduction with either caffeate or ferulate. *p*-Coumaric-2-^{13}C, caffeic-2-^{13}C and ferulic-2-^{13}C acid were synthesized from the corresponding benzaldehyde and 2-^{13}C-malonic acid (cat. piperidine in pyridine, 68 ^{o}C oil bath, 24 h) and fed to cassava roots. All three intermediates were incorporated into labelled scopoletin and its glycoside scopolin. On average, 21% of scopoletin was labelled. We conclude that the major pathway for scopoletin and scopolin biosynthesis in cassava during PPD is through *trans*-2',4'-dihydroxycinnamate, enzymatically isomerised and then ring closed to umbelliferone, but all the three pathways are present and functioning to varying extents. These are being investigated with ^{18}O-isotopic enrichment of umbelliferone (0.5 ml 95% labelled $H_2{}^{18}O$, 10 µl conc. HCl, 1 ml MeCN, 7 days, 70 ^{o}C oil bath) which gave 82% of labelled umbelliferone, $C_9H_7{}^{16}O_2{}^{18}O$ HR ESI-MS found *m/z* 165.0430 $[M + H]^+$ (calcd. 165.0432) and *trans*-cinnamic acid (0.5 ml 95% labelled $H_2{}^{18}O$, conditions as above, 4 days) which gave 86% of double labelled *trans*-cinnamic acid, $C_9H_8{}^{18}O_2$ HR ESI-MS found *m/z* 153.0687 $[M + H]^+$ (calcd. 153.0682) which is being used to investigate *ortho*-hydroxylation or alternatively a spirodienone intermediate leading to cyclisation in scopoletin biosynthesis in PPD.

We acknowledge the financial support of the Egyptian Government (studentship to SALB).

1. F. Bourgaud, A. Hehn, R. Larbat, S. Doerper, E. Gontier, S. Kellner, and U. Matern, *Phytochem. Rev.*, 2006, 5, 293-308.
2. K. Reilly, R. Gomez-Vasquez, H. Buschmann, J. Tohme, and J.R. Beeching, *Plant Mol. Biol.*, 2003, 53, 669-685.
3. K. Kai, B. Shimizu, M. Mizutani, K. Watanabe, and K. Sakata *Phytochem.*, 2006, **67**, 379-386.

Investigation of scopoletin biosynthesis during post-harvest physiological deterioration in cassava roots using stable isotopic labelling

Soad A. L. Bayoumi, Michael G. Rowan, Ian S. Blagbrough, and John R. Beeching*
Department of Pharmacy and Pharmacology, University of Bath, Bath BA2 7AY, and
*Department of Biology and Biochemistry, University of Bath, Bath BA2 7AY, U.K.
E-mail: salhb20@bath.ac.uk, prsmgr@bath.ac.uk, prsisb@bath.ac.uk, bssjrb@bath.ac.uk

Objectives Coumarins are pharmacologically active and have roles in plant defence. Despite their importance, key aspects of the biosynthesis of these secondary metabolites remain unresolved. Here we exploit the observation that the accumulation of scopoletin and its glucoside scopolin increases in cassava roots (*Manihot esculenta* Crantz) during post-harvest physiological deterioration (PPD) to test alternative pathways for the biosynthesis of these hydroxycoumarins.

Methods Cassava roots (cv Mcol22 and NGA 19), within 2 h of harvesting, were fed with different labelled intermediates on the postulated biosynthetic pathway of scopoletin (e.g. *trans*-cinnamic-d_7, *trans*-cinnamic-3,2',3',4',5',6'-d_6, *trans*-cinnamic-2',3',4',5',6'-d_5, *trans*-cinnamic-2-d, *trans*-cinnamic-$^{18}O_2$, *p*-coumaric-2-^{13}C, caffeic-2-^{13}C and ferulic-2-^{13}C acids. Also, competition feeding experiments with a mixture of *trans*-cinnamic-2',3',4',5',6'-d_5 with each of the following: 2',4'-dihydroxycinnamic, caffeic, or ferulic acids) were carried out. PPD was allowed to occur. Ethanolic extracts of the deteriorated roots were separated (HPLC) and analysed (HR ESI-MS).

Results Deuteriated cinnamic acids were incorporated, and typically 29% of the scopoletin was deuteriated. Incorporation (in both scopoletin and scopolin) of only 3 deuterons when fed with *trans*-cinnamic-d_7, *trans*-cinnamic-d_6 and non deuteriated scopoletin when cinnamic-2-d acid was fed indicates that the pathway involves exchange of the 2-hydrogen atom in cinnamic acid, and strongly supports our thesis that the *trans*-*cis*-isomerisation step is enzymatic, and not photochemical, as found *in vitro* and in other plants where 4 deuterons would have been found in the labelled product. Incorporation of *p*-coumaric-2-^{13}C, caffeic-2-^{13}C and ferulic-2-^{13}C acids in the biosynthesis of scopoletin gave an average increase of 14% ^{13}C-labelled scopoletin and ^{13}C-labelled scopolin. There was no reduction in label incorporation when either caffeic or ferulic acids were fed in competition to deuterium-labelled cinnamic acid whereas competition with unlabelled *trans*-2',4'-dihydroxycinnamate caused a decrease from 4.6 to 3.4% (compared to the labelled scopoletin isolated when the roots were fed with labelled cinnamic acid only). *Trans*-cinnamic-$^{18}O_2$ was incorporated (5%) in the biosynthesis of scopoletin with only one ^{18}O-labelled oxygen atom in the product.

Conclusions The ready accumulation of scopoletin and scopolin in cassava roots during PPD make it a good model to investigate their biosynthesis. We have shown that the *E*-*Z*-isomerisation step (of the cinnamic acid derivative) during the biosynthesis is enzymatic. Furthermore, all three proposed pathways (as found in different plants) are operating in cassava, but the pathway via 2',4'-dihydroxycinnamate is likely to be the predominant one. A pathway via a spirolactone-dienone (quinol) intermediate has been previously established in *Streptomyces niveus* for novobiocin biosynthesis in elegant work by Kenner and co-workers (Bunton et al 1963), and also proposed from UV studies in cultures of the plant *Ammi majus* L. (Apiaceae, Bishop's flower, large bullwort) (Matern 1991) following work by Grisebach and Ollis (1961). The absence of doubly enriched ^{18}O-scopoletin means that the lactonisation step is through *ortho*-hydroxylation not via a spirolactone-dienone intermediate where both ^{18}O-atoms would be incorporated in the final product.

We acknowledge the financial support of the Egyptian Government (studentship to SALB).

References

Bunton, C. A., et al (1963) *Tetrahedron* **19**: 1001-1010.
Grisebach, H., Ollis, W. D. (1961) *Experientia* **17**: 4-12.
Matern, U. (1991) *Planta Med*. **57**: S15-S20.